Yeast adaptation and survival under acute exposure to lethal ethanol stress

Jamie S. Yang

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

under the Executive Committee

of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Yeast adaptation and survival under acute exposure to lethal ethanol stress

Jamie S. Yang

The ability to respond to stress is universal in all domains of life. Failure to properly execute the stress response compromises the fitness of the organism. Several key stress pathways are conserved from unicellular organisms to higher eukaryotes, so knowledge of how these pathways operate in model organisms is crucial for understanding stress-related diseases and aging in humans. The mechanisms of stress tolerance have been well-studied in the budding yeast *Saccharomyces cerevisiae*. Yeast respond to diverse stresses by initiating both general and stress-specific responses that generally protect the cells during and after the stress exposure. While previous work has revealed mechanistic insights on adaptation and survival under mild and long-term exposure to stress, how they cope with acute exposure to lethal stress is not well understood.

Here, we combined transcriptional profiling, fitness profiling, and laboratory evolution to investigate how *S. cerevisiae* survive acute exposure to lethal ethanol stress. By using high throughput methods such as RNA-seq and barcode sequencing of the pooled yeast deletion library, we were able to discover and characterize both existing and novel pathways that yeast utilize to adapt to and survive ethanol stress. We found both ethanol-specific and as well general stress response mechanisms. We were also able to evolve a strain of ethanol under lethal ethanol stress to exhibit a survival of at least an order of magnitude greater than the parental wild-type strain. Additionally, this evolved strain exhibited cross protection to other stresses without compromising bulk growth rate. We found that this strain adapted its global expression levels to a post-stress state, making it more robust to various stresses even under optimal growth conditions.

Table of Contents

List of Figures and Tables

Acknowledgments

First and foremost, I would like to express my utmost appreciation for the mentorship of my thesis advisor Saeed Tavazoie. In addition to his brilliance as a scientist, he has also been an exceptionally supportive mentor. The quality of his alumni is a testament to his ability in fostering one of the most collaborative and supportive laboratory environments I have ever encountered. I am deeply grateful for his endless support and guidance, allowing me to grow significantly as a scientist.

I cannot thank enough all the members of the Tavazoie Lab for their unwavering support and generosity in helping me with all aspects of my training as a graduate student. I could not have asked for a better group of postdoctoral mentors: Panos Oikonomou, Wenyan Jiang, Amir Momen Roknabadi, Balaji Santhanam, Anu Khare, and Cindy Fang, all of whom have helped me so much in various aspects of my project. I am especially grateful to a former postdoctoral researcher in the lab, Peter Freddolino, whom I had the pleasure of working closely with for the first few years as a graduate student. His mentorship and guidance have been indispensable, always there for me whenever I needed any support. I am very thankful to have the support of fellow graduate students Alexandra Ketcham and Sydney Blattman, who continue to motivate and inspire me with their own wonderful work.

I am indebted to my excellent thesis committee, Lorraine Symington, Sagi Shapira, and Harris Wang, for their wonderful insights and invaluable advice. Additionally, I would like to thank Marko Jovanovic as my outside examiner on my defense committee.

I would also like to give a special thank you to Zaia Sivo, who has been there to guide me every step of the way of my journey. Additionally, I am thankful for all the support and mentorship of the MD/PhD directors, Michael Shelanski, Steven Reiner, Patrice Spitalnik, and Ronald Liem.

My friends and family have been an essential source of support and happiness throughout my graduate years. My sister Lena has always been there for me through any ups and downs during these few years. My wonderful parents have provided me with much wisdom, encouragement, and unconditional love. I am incredibly lucky to be a brother and son to them. Lastly, I would like to thank my loving wife Alice for always being by my side through all the late nights and moments when I was stuck, especially for feeding me through all the innumerable times I decided to choose science over food.

DEDICATION

To the memory of my grandmother, Fang Shaofeng (1934-2016), for her love and care.

INTRODUCTION

Organisms in the wild are constantly being exposed to various environmental changes. They will often sense and properly respond to these alterations in their surroundings to maximize their fitness in a given environment. Some environmental changes are positive, such as the availability of a food source. Organisms may respond to that change by moving towards the food source or increasing uptake of the nutrient. Other environmental changes may be negative. These perturbations threaten the organism in some way and may cause cellular or macromolecular damage to the orgasm. We refer to these negative environmental changes as stress. Many stresses work in a dose-dependent manner, with higher stress levels causing more damage to the orgasm. Some stress levels are mild and will not kill an orgasm in its wild-type state. We refer to these stresses as sublethal. As the level of stress increases, there may be a point at which some members of a wild-type population start to die. This is the threshold of lethality. Above this level of stress, a greater fraction of the population will die, and the stress at this level is considered lethal. While stresses at a high enough dose will be lethal, organisms have evolved mechanisms to sense and adapt to mild levels of stress. Cellular responses to many stresses have been widely described in model organisms, especially in *Saccharomyces cerevisiae*, and a study of the yeast stress response helps elucidate mechanisms and pathways of how cells survive and adapt to stress.

Stress response in yeast

A particularly well studied example of the yeast stress response is its response to heat shock (Lindquist and Craig, 1988; Mager and Ferreira, 1993; Piper, 1993; Mager and De Kruijff, 1995). The heat shock proteins that were discovered to mediate this response were found to be induced in other stresses as well. Several studies also found that mild heat shock protected yeast cells from not only subsequent exposure to lethal levels of heat shock, but also exposure to lethal levels of

other stresses, such as desiccation stress (Hottiger et al., 1987), oxidative stress (Benaroudj et al., 2001), and ethanol stress (Sanchez et al., 1992). This biological phenomenon was known as cross protection, in which yeast cells exposed to a mild stress develop tolerance to other stresses. These observations led stress biologists to propose the existence of a general stress response, in which different stresses create the same intracellular signal (Ruis and Schüller, 1995). A general stress transcription factor was proposed when genes that were found to be upregulated or downregulated under one stress condition were similarly up- or downregulated under different stress conditions (Kobayashi and McEntee, 1990; Wieser et al., 1991). The promoters of these stress-responsive genes contained a common sequence, with a core consensus sequence of AGGGG, known as the stress response element (Kobayashi and McEntee, 1993; Marchler et al., 1993; Schüller et al., 1994).

The transcription factors that bind the stress response element were found to be zinc finger proteins called *MSN2* and *MSN4*, and they were coined as the general stress transcriptional factors (Martínez-Pastor et al., 1996). A *ΔMSN2 ΔMSN4* double mutant was sensitive to various different stresses, and overexpression of the two genes improved resistance to some of the same stresses. Pathways that were important for growth in yeast were found to repress expression of genes dependent on *MSN2* and *MSN4*. This included the protein kinase A (PKA) pathway (Smith et al., 1998) and the target of rapamycin (TOR) pathway (Beck and Hall, 1999). In the absence of *MSN2* and *MSN4*, however, some of their targets may still be induced in response to stress, hinting at a more complicated system of stress regulation. For example, the transcription factor *YAP1*, discovered to be a key mediator of oxidative stress, could also be activated by a variety of stresses and had a general stress component (Kuge et al., 1997; Wemmie et al., 1997).

Environmental stress response

The development of global analyses allowed researchers to more fully understand the responses to environmental stresses. Exploring genomic expression patterns responding to a diverse set of stresses using DNA microarrays allowed researchers to implicate a significant number of previously uncharacterized genes in the general stress response (Gasch et al., 2000; Causton et al., 2001). These two studies covered genomic expression studies characterizing wild-type yeast in response to heat shock, pH extremes (acid, alkali), oxidative and reductive stress (hydrogen peroxide, menadione, diamide, dithiothreitol), hyper-osmotic shock (sorbitol, potassium chloride, sodium chloride), and starvation (progression into stationary phase, amino acid starvation, nitrogen starvation). Other studies around the time also performed genomic expression studies involving ethanol shock (Alexandre et al., 2001), cadmium exposure (Momose and Iwahashi, 2001), phosphate starvation (Ogawa et al., 2000), zinc starvation (Lyons et al., 2000), copper starvation (Gross et al., 2000), diverse drug treatments (Bammert and Fostel, 2000; Hughes et al., 2000), anaerobic growth (Kwast et al., 2002), ionizing radiation (Sanctis et al., 2001; Gasch et al., 2001), and double-strand breaks (Lee et al., 2000). A large set of genes showed a similar response to the majority of these stresses.

This global expression pattern in response to these diverse stresses was coined by Gasch et al., 2000 as the environmental stress response (ESR). This response consisted of a set of approximately 900 genes, about 600 of which were repressed and about 300 of which were induced by these various stresses. The expression pattern of the induced and repressed groups of genes show almost identical but opposite patterns of expression in response to various stresses (Figure 1.1). The magnitude of expression changes in the ESR is proportional to the degree of stress

sensed by the cell. The transcription factors implicated in this response were *MSN2*, *MSN4*, and *YAP1*, same as those discovered for the general stress response.

Two groups with distinct expression profiles were found within the repressed genes of the ESR. One group of genes appear to be coregulated and have functions involved in growth-related processes, RNA processing and splicing, translation initiation and elongation, tRNA synthesis and processing, nucleotide biosynthesis, secretion, and other metabolic processes. The second group has a slight delay in the downregulation of expression and is mainly composed of genes that encode ribosomal proteins. Repression of genes encoding ribosomal proteins, RNA metabolism, protein synthesis, and cell other aspects of cell growth is a defining feature of the repressed ESR (Gasch et al., 2000). The reduced synthesis of genes required for ribosome biogenesis, ribosomal proteins, and translation initiation was proposed to conserve energy while the cell adapts to the new stressful environment (Warner, 1999; Gasch et al., 2002).

Regarding the group of genes induced in the ESR, many of those genes were previously known to be protective of the cell during stress conditions. A wide variety of processes were found in the induced ESR, such as carbohydrate metabolism, detoxification of reactive oxygen species, cellular redox, cell wall modification, protein folding and degradation, DNA damage repair, fatty acid metabolism, metabolite transport, autophagy, intracellular signaling, vacuolar functions, and mitochondrial functions. Many of the induced ESR genes also played reciprocal functions, such as the upregulation of genes that code for synthesis of enzymes and also induction of genes that code for catabolism of enzymes. The same scenario existed for genes regulating the protein kinase A pathway. It was proposed that this observation allowed the cells finer control of these specific pathways by shifting the regulation to be controlled translationally (Gasch et al., 2000).

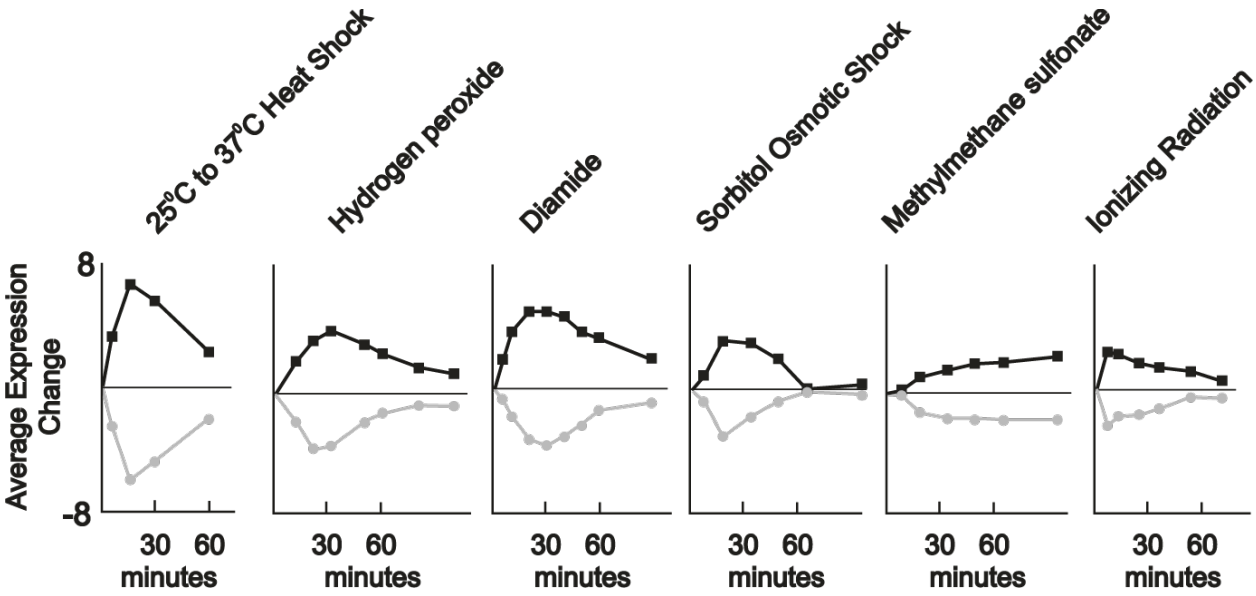


Figure 1.1 The environmental stress response
The average gene expression levels of ESR genes whose expression is induced (black lines) or repressed (gray lines) in response to six different mild stresses. The stress levels are as follows: 25°C to 37°C heat shock, 0.3 mM hydrogen peroxide, 1.5 mM diamide, 1 M sorbitol, 0.02% methyl methanesulfonate, and 170 Gray of ionizing radiation. Figure from Gasch, 2003.

While the ESR is a common response to many different stresses, its regulation is more specific to the particular stress, being controlled at many levels, such as transcriptional initiation, chromatin remodeling, and mRNA turnover (Gasch et al., 2002). For example, the transcription factors *HSF1*, *HOT1*, and *YAP1* independently affect different subsets of the ESR in response to heat shock, osmotic stress, or oxidative stress, respectively (Treger et al., 1998; Rep et al., 1999, 2000; Amoros and Estruch, 2001). This indicates that the cell can activate the ESR in response to a wide-variety of upstream signals.

Growth rate versus stress

Since the repressed ESR included many genes related to growth-related processes, researchers wondered if the ESR was actually a growth rate response. More than 70% of the genes that were repressed as part of ESR were involved in protein synthesis, ribosome synthesis and

processing, RNA polymerase I- and III-dependent transcription, and translation, all processes that affect cellular growth (Ashburner et al., 2000; Ball et al., 2000). Genome-wide expression studies were performed in which growth rates were controlled in chemostats by limiting nutrients such as glucose, ammonium, phosphate, sulfate, leucine, or uracil (Castrillo et al., 2007; Brauer et al., 2008). It was found that expression of more than one fourth of all yeast genes was linearly correlated with growth rate, independent of the nutrient that was being limited. Up to one third of all yeast genes could be correlated with growth rate for certain limiting nutrients. Additionally, positively correlated genes with growth rate mainly encoded for ribosomal functions (Brauer et al., 2008). When ESR genes were examined for their correlation to growth rate, it was found that repressed ESR genes were positively correlated with growth rate, and induced ESR genes were negatively correlated with growth rate. Lu et al., 2009 suggested two possibilities that could explain this observation. The first was that stress directly causes slow or stopped growth, and this growth rate reduction would cause the gene expression changes seen in the ESR. The second possibility was that slow growth itself was perceived by the cells as stress. They performed experiments to disentangle the direct effects of stress from the changes in growth rate (Lu et al., 2009).

Using a chemostat, yeast cultures were grown under phosphate limitation at four different growth rates ranging from a 0.072 doublings per hour to 0.43 doublings per hour. A sample from each culture was then acutely heat shocked to a temperature of 50℃. They discovered that the fastest growing cells were the most sensitive to heat shock, the slowest growing cells were resistant to heat shock, and cells growing at intermediate rates showed intermediate rates of killing, indicating a negative correlation between growth rate and survival under heat shock. Performing the same heat shock experiments using different nitrogen sources to control for growth rate in

batch culture confirmed this negative correlation between growth rate and stress defense as well (Lu et al., 2009).

Another approach to study the connection between growth rate and stress defense is by using the yeast deletion collection, which contains deletions of about 4,800 haploid deletions out of the more than 6,000 possible ORFs (http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3.html). Many gene deletion strains show a slow growth phenotype when grown in rich media at 30℃. Their growth defects ranged from 12% to 90% of the wild-type growth (Giaever et al., 2002). The genes that were required for optimal growth were enriched in the categories of protein synthesis, ribosomal proteins, mitochondrial functions, and respiration (Mewes et al., 2000). The yeast deletion library can be used to systematically investigate the genes and pathways involved in stress survival. Additionally, all deletion strains of the same background can be pooled together to simultaneously test the effects of each gene deletion.

Zakrzewska et al., 2011 used the pooled yeast deletion library to test stress survival. The pooled library was grown for three hours at either optimal growth conditions of 30℃ or mild heat stress at 38℃. The populations were exposed to ten minutes of three different lethal stresses, consisting of either hydrogen peroxide, acetic acid, or heat. The samples were then recovered and DNA was isolated for hybridization to fitness arrays (Zakrzewska et al., 2011). For cells that were not pre-exposed to mild heat shock, the growth rate of the mutants was a strong determinant of stress survival, with up to half of the differences in severe stress survival being accounted for by differences in growth rate (Figure 1.2). In contrast, only 10% - 25% of the results could be explained by mechanistic overlap.

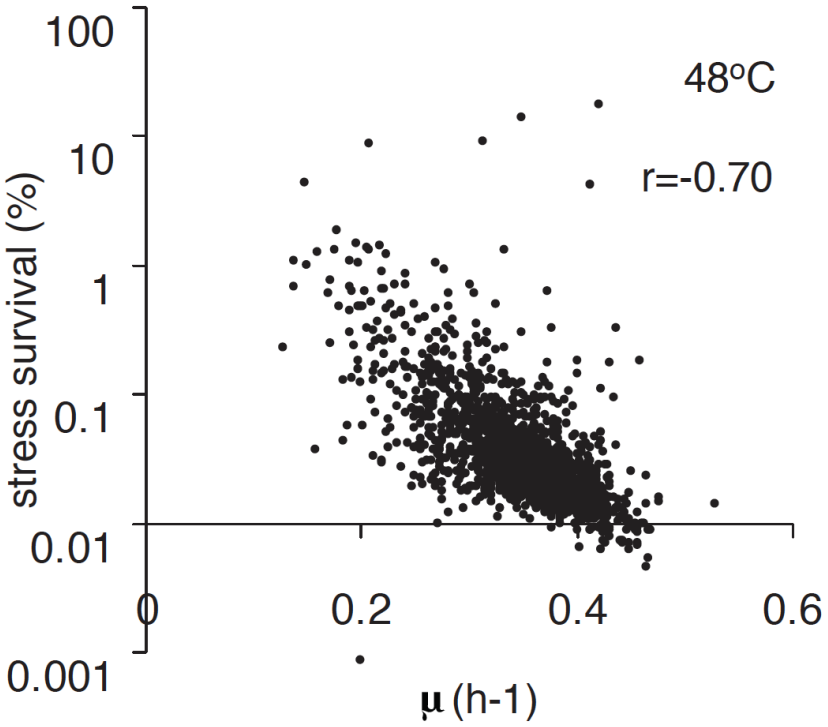


Figure 1.2 Inverse correlation between stress survival and growth rate
The survival percentage of each gene deletion strain after exposure to 10 minutes of 48°C heat shock is plotted against each of their growth rates in rich media. Figure from Zakrzewska et al., 2011.

To determine the pathways that were important for stress survival, the contribution of growth rate to severe stress survival was corrected for using linear regression. The deletion strains that survived better due to growth rate were enriched for functions in RNA polymerase II mediated transcription, translation, ribosome biogenesis, mitochondrial functions, and DNA and chromosome organization. After growth rate correction, stress-specific pathways were able to be identified. Additionally, the correlation between any two individual stresses decreased after growth rate correction, confirming that cross resistance was determined mostly by growth rate (Zakrzewska et al., 2011).

The previous two studies showed an inverse correlation between severe stress survival and growth rate by altering growth rate either through nutrient limitation or through gene deletions. Nutrient limitation and certain gene deletions can, however, cause increased survival through

mechanisms other than purely growth rate. Using wild-type yeast that have a range of growth rates would be more ideal to test the effects of growth rate without confounding variables. Levy et al., 2012 used a high-throughput microscopy growth assay to measure the growth rate variance of wild-type yeast. Using time-lapse images, they were able to determine the growth rate of different colonies on a plate. They found that wild-type populations growing in optimal conditions on a plate had about 1.3% - 10% of their colonies growing at less than half the median population growth rate, a trait that was also heritable. They ruled out all genetic mechanisms and mitochondrial aberrations for explaining the slow growth, describing it instead as a bet-hedging mechanism. The stress experiments were performed by growing the cells under optimal growth conditions for six hours and then heat shocking the cells so that most cells are killed. The cells were observed under the microscope from the start of growth until 14 – 20 hours after the heat shock. All colonies that did not have a slow growth rate died from the heat shock. Many slow growing cells, however, survived the heat shock, underwent one to two cell divisions at a slow rate, and then produced fast growing progeny (Levy et al., 2012).

This study also used a marker, Tsl1, which had a negative correlation with growth rate. Using fluorescence activated cell sorting on Tsl1-GFP cells, they were able to sort for cells into bins based on Tsl1 levels, which was a proxy for growth rate. They found a dose-dependent effect between Tsl1 and stress survival. Higher Tsl1 levels predicted higher chance of stress survival. They concluded from this study that populations of yeast contain a continuous range of metastable epigenetic states that each confer a different fitness. The states in which the cells are slowest growing protect against stresses, whereas the states in which the cells are fastest growing are the most susceptible to stress (Levy et al., 2012).

The association of stress survival and growth rate was so strongly linked that many believed that rapid growth and stress defense were competing interests for the cell, claiming that is it not possible to optimize one independently of the other. The known pathways that drive growth in yeast (TOR and RAS/PKA pathways) repress stress defense, and vice versa (Figure 1.3).

Lethal stress in yeast

The majority of studies on yeast stress survival have focused on acquiring stress tolerance after exposing yeast to a mild dose of stress so that they can better survive subsequent doses of lethal stress. Few studies have looked at acute exposure to transient lethal stresses only without any pre-exposure to mild doses of stress. Many studies have investigated nutrient deprivation and stationary phase, which are also lethal stresses, but we will not discuss those further because they are a completely different type of stress. They are stresses in which the cell can never overcome unless the nutrient is added back into the cell's media. Therefore, experiments can only be performed to increase their lifespan or decrease their rate of death in that media rather than overcome the stress and proliferate again (Longo et al., 2012). Additionally, because cells die over a long period of time during nutrient deprivation or stationary phase, there is much time for cells to mount a response to slow down their rate of death or even enter a dormant state that is highly stress-resistant. We are interested in lethal stresses in an acute setting that minimizes the time the cells can mount a response to the stress.

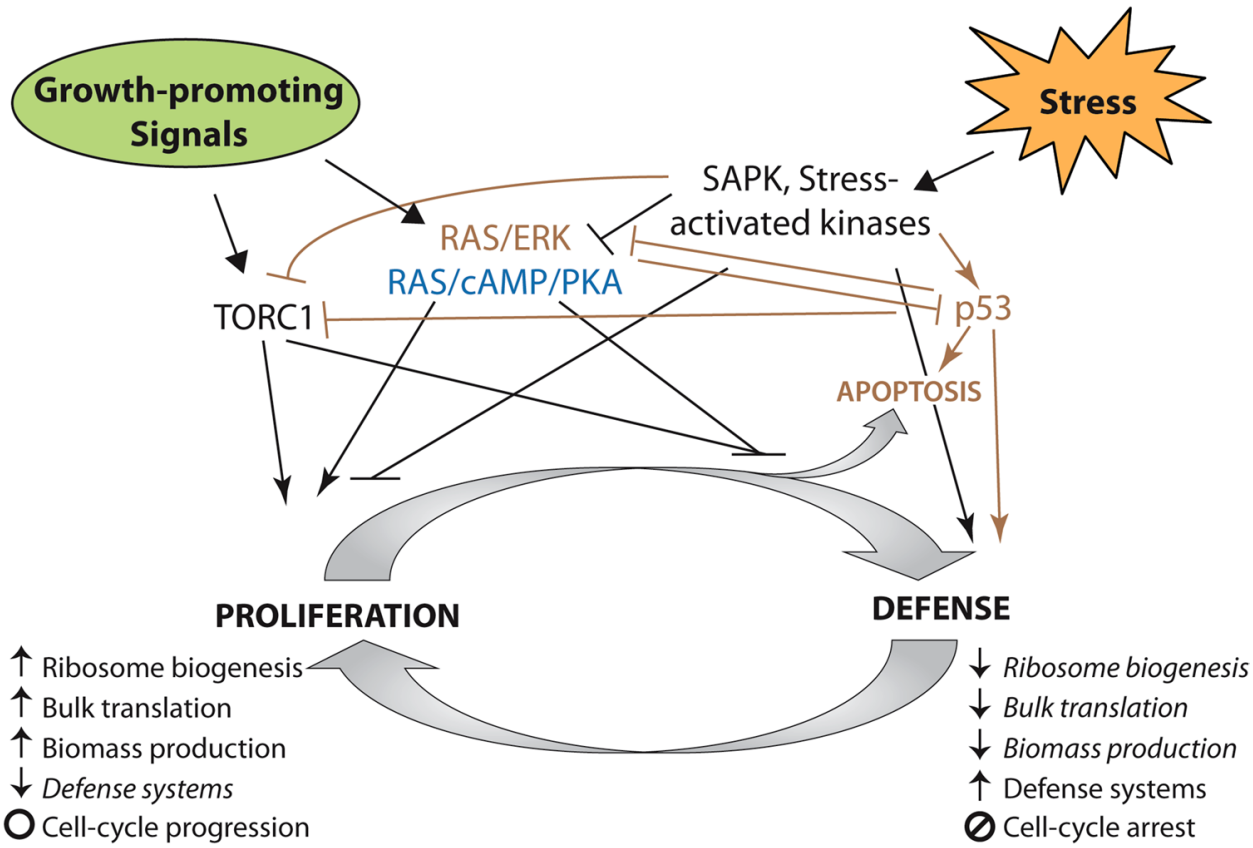


Figure 1.3 Growth and stress compete for limited cellular resources
The signaling processes that occur in yeast and mammalian cells are shown here. Pathways relevant to yeast are shown in blue, pathways relevant to mammalian cells are shown in brown, and pathways common to both are shown in black. Figure from Ho and Gasch, 2015.

As mentioned earlier, Zakrzewska et al., 2011 used the yeast deletion library to assess severe stress survival to three different stresses. While the majority of the study focused on the acquisition of stress tolerance following a mild heat shock, they did list the functional groups that were responsible for severe stress survival without a pre-exposure to any mild stress. After correcting for growth rate, they found that severe oxidative stress enriched for gene deletions with functions in 90S pre-ribosome, ribonucleoprotein complex, rRNA processing, ribosomal protein, RNA elongation, tRNA wobble uridine modification, and ribosome biogenesis. Severe heat shock enriched for gene deletions in mitochondrial functions, translation, RNA polymerase II mediated transcription, RNA elongation, protein targeting to vacuole, stress granule, chromatin

modification, and chromosome organization (Zakrzewska et al., 2011). From these results, one can see that survival to extreme oxidative stress is helped by deletions in RNA processing, ribosomal proteins, and ribosome biogenesis, all functions that are correlated with growth rate. Similarly, survival to extreme heat shock is helped by deletions in mitochondrial functions, translation, and RNA polymerase II mediated transcription, all functional groups that are also correlated with growth rate. This suggests that these functions have an effect on lethal stress survival that is more than simply based on growth rate. It is possible that in addition to affecting cellular growth, these processes also actively divert cellular resources towards stress defense.

Another extreme stress survival study, Yaakov et al., 2017, looked at exposure to lethal levels of the antifungal fluphenazine. They found that cells with extremely high levels of heat shock protein 12 (Hsp12) had increased survival to fluphenazine. They discovered that Hsp12 also correlated with DNA damage and hypothesized that DNA damage was the cause of the increased fluphenazine survival. They were able to show that random double-stranded breaks in a small population of a yeast population activated the ESR internally, causing those cells to be more stress resistant. Creating the mutant *ΔMSN2 ΔMSN4* decreased their survival to fluphenazine, as expected, but also decreased the percentage of cells with extremely high levels of Hsp12. In addition to the ESR activation, increased double-stranded breaks in a small population predisposes those cells to have more mutations, which allows the cells to have a greater chance of acquiring a mutation that is beneficial under stress (Yaakov et al., 2017). This is similar to Levy et al., 2012 in that it also suggested a bet hedging strategy that allows a small fraction of population to be highly stress resistant, which increases the fitness of the population under environments that are prone to be stressful.

Yeast response to ethanol stress

The study of the yeast response to ethanol stress has been of particular interest in the food industry, for making beer and wine, and also in the biofuel industry, for making ethanol. High ethanol concentrations, however, will inhibit cell growth and limit ethanol yield and productivity from fermentation (Ansanay-Galeote et al., 2001; Aguilera et al., 2006). Understanding the ethanol stress response and improving survival to high alcohol concentrations can significantly improve cost-effective ethanol production in these industries.

The main sites of ethanol action in yeast are the plasma membrane, hydrophobic proteins of the cell and mitochondrial membranes, nuclear membrane, vacuolar membrane, endoplasmic reticulum, and hydrophilic proteins in the cytoplasm (D'amore et al., 1989; Walker, 1998). By interacting with these membranes, ethanol increases membrane fluidity and decreases membrane structural integrity (Mishra and Prasad, 1989). Vacuolar morphology is also altered under ethanol stress in that segregated structures become one large organelle (Meaden et al., 1999). Important enzymes and proteins can be denatured due to the decrease in cellular water content (Hallsworth et al., 1998). Additionally, ethanol induces the production of heat shock-like proteins, lowers the rate of RNA and protein accumulation, and increases the amount of petite mutations (Huo et al., 2006). The main effects of yeast exposure to ethanol are shown in Table 1.1.

Cell function and ethanol influence	Source
Cell viability and growth	
Inhibition of growth, cell division and cell viability	Stanley *et al.* (1997)
Decrease in cell volume	Birch and Walker (2000)
Metabolism	
Lowered mRNA and protein levels	Chandler *et al.* (2004), Hu *et al.* (2007)
Protein denaturation and reduced glycolytic enzyme activity	Hallsworth *et al.* (1998)
Induction of heat shock proteins and other stress response proteins	Plesset *et al.* (1982)
Intracellular trehalose accumulation	Lucero *et al.* (2000)
Cell structure and membrane function	
Altered vacuole morphology	Meaden *et al.* (1999)
Inhibition of endocytosis	Lucero *et al.* (2000)
Increased unsaturated/saturated fatty acid ratio in membranes	Alexandre *et al.* (1994)
Increase in ergosterol content of membranes	Sajbidor *et al.* (1995)
Loss of electrochemical gradients and proton-motive force	Petrov and Okorokov (1990)
Inhibition of transport processes	Leao and van Uden (1984)
Inhibition of H^+-ATPase activity	Cartwright *et al.* (1986)
Increased membrane fluidity	Mishra and Prasad (1989)

Table 1.1 Effects of ethanol stress on yeast
Figure from Stanley et al., 2009.

Several transcriptome-wide studies were performed comparing the transcriptomes of stressed versus non-stressed yeast cells during short term sublethal ethanol exposure. The gene ontology categories that were found to be upregulated upon exposure to ethanol stress were cell energetics, transport mechanisms, cell surface interactions, lipid metabolism, general stress response, trehalose metabolism, protein destination, ionic homeostasis, glycolysis, and

tricarboxylic acid cycle (Alexandre et al., 2001; Chandler et al., 2004; Fujita et al., 2004). There was a significant increase in genes encoding for hexose transport and glycolysis, which led Chandler et al., 2004 to propose that yeast enter a pseudo-starvation state during ethanol stress. The genes that decreased in expression upon exposure to ethanol stress contained gene ontology categories of protein synthesis, RNA synthesis and processing, amino acid metabolism, and nucleotide metabolism (Alexandre et al., 2001; Chandler et al., 2004). Chandler et al., 2004 also found that late-stage ethanol response genes (3 hours post-stress) only shared 7% commonly with early-stage response genes (one hour post-stress).

A number of studies also used the pooled yeast deletion library to determine gene deletions that were sensitive or tolerant to ethanol (Kubota et al., 2004; Fujita et al., 2006; Van Voorst et al., 2006; Yoshikawa et al., 2009). The severity of ethanol stress was found to influence which genes were sensitive to ethanol. Different genes were associated with ethanol tolerance when exposed to 8% versus 11% ethanol. At 11% ethanol, the genes necessary for ethanol tolerance had functions in biosynthesis, cell cycle, cytoskeleton, mitochondria, morphogenesis, nucleic acid binding, protease activity, protein transport, vacuole, signal transduction, transcription, and transport (Kubota et al., 2004). At 8% ethanol, the main functional categories associated with ethanol-sensitive deletion strains were tryptophan metabolism, vesicular and vacuolar transport, mitochondrial functions, and peroxisomal transport (Yoshikawa et al., 2009). If all four of the pooled yeast deletion library results are combined, there was not a large overlap among the genes in each study, possibly indicating difference in strains and ethanol concentrations used in the studies.

While many of the genes and pathways from the previous studies is associated with mechanisms of the general stress response, some ethanol-specific responses have also been

discovered. While both ethanol and heat stress cause selective mRNA export, bulk poly(A)+ mRNA accumulates in the nucleus under ethanol stress. Nuclear localization of the DEAD box protein Rat8 in response to ethanol stress blocked bulk poly(A)+ mRNA export, but there was not a similar Rat8 localization under heat stress, suggesting specificity of Rat8 to ethanol stress (Takemura et al., 2004). Asr1 was also identified as an ethanol-specific transcriptional regulator. It is a yeast Ring/PHD finger protein that normally shuttles between the cytoplasm and nucleus constitutively, but upon ethanol stress, it rapidly accumulates in the nucleus. This protein does not show similar localization under oxidative, osmotic, nutrient limitation, or heat stress (Betz et al., 2004). The authors suggested that Asr1 might allow yeast to acclimate to ethanol. A later study showed, however, that there was no significant difference between *ΔASR1* and wild-type strains, suggesting instead that Asr1 accumulation in the nucleus was only due to a failure of nuclear export machinery under ethanol stress (Izawa et al, 2006).

METHODS

Media and growth conditions

We used YPD broth (10 g/L Bacto yeast extract, 20 g/L Bacto peptone, 20 g/L glucose) and YPD agar plates (YPD broth with 20 g/L Bacto agar) for routine growth of yeast strains. All experiments were performed using standard complete + glucose (sc+glu) media (6.7 g/L yeast nitrogen base (Difco), 2 g/L yeast synthetic drop-out mix (US Biologicals), 20 g/L glucose) and sc+glu plates (sc+glu media with 20 g/L Bacto agar). All growth on plates occurred at 30℃. All growth in liquid media also occurred at 30℃ and shaking at 220 rpm in an Innova 42 incubator (New Brunswick). To test growth rate, the number of cells were tracked by measuring optical density at a wavelength of 660 nm using an Ultrospec 3100 pro spectrophotometer (Biochrom).

Yeast strains

All yeast strains were derived from BY4741 (Mat **a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) (Brachmann et al., 1998), which is the strain we refer to as wild-type. All deletion strains were obtained from the *S. cerevisiae* knockout collection (Giaever et al., 2002). The gene deletions used in this study are *ΔTOR1* and *ΔTCO89*.

Stress experiments

All stress experiments, regardless of background or using the pooled yeast deletion library, were performed identically. Cells were first grown to mid-log phase in 25 mL sc+glu in a 250-mL flask at 30℃ and then washed. They were transferred to 50 mL Falcon tubes and centrifuged for five minutes at 3000 rcf to pellet in an Allegra 25R centrifuge (Beckman Coulter). The supernatant was discarded and the cells were resuspended in 25 mL deionized water. They were again centrifuged for five minutes at 3000 rcf to pellet. The supernatant was discarded and the cells were

resuspended in 1 mL of deionized water. They were then centrifuged at 20,000 rcf to pellet on a benchtop centrifuge. The supernatant was discarded and the cells were resuspended in 600 microliters of deionized water. The pre-stress time points for all experiments were taken from this cell suspension.

To stress the cells with ethanol or hydrogen peroxide, 400 microliters of the suspension was added to 600 microliters of solution containing the appropriate stress, immediately vortexed, set at room temperature for two minutes, vortexed again, and immediately removed from the stress. The 600 microliters of solution contained the amount of stress such that the final 1 mL of cell suspension contained the desired concentration of ethanol or hydrogen peroxide. For example, if the cells were to experience a stress level of 20% ethanol, 200 microliters of 100% ethanol were added to 400 microliters of deionized water. As soon as 400 microliters of cells were added to this solution, the final 1 mL volume of cells would immediately encounter the 20% ethanol stress. In the case of heat stress, 400 microliters of the pre-stress cell suspension were added to 600 microliters of deionized water. This 1 mL of cell suspension was vortexed, placed in a water bath at the desired temperature for two minutes, vortex again, and immediately removed from the stress.

For all three types of stress used, the stress was removed through dilution. The 1 mL of stressed cells was diluted tenfold into deionized water and vortexed. The post-stress time points for all experiments were taken from this final cell suspension, whether it be for recovery or plating. A control experiment was performed in which cells were subjected to the tenfold diluted concentration of the desired stress. They were in this mild stress for four hours and plated periodically on sc+glu plates. There was no loss in viability in any of the tenfold diluted concentrations of either hydrogen peroxide or ethanol stress.

For all experiments in which fraction survival was calculated, the 100 microliters were removed from the pre-stress and post-stress cell suspension for serial dilutions in deionized water and then plated on sc+glu plates. The number of colonies (CFU) from the plates were counted daily until the CFU count no longer increased. The number of cells before and after stress were calculated from the CFU counts. Fraction survival was defined as the ratio of post-stress cells to pre-stress cells.

Experimental setup for transcriptional profiling using RNA-seq

Diagrammatically shown in Figure 3.4, there were five time points taken for each experiment, which will be referred to as growth, pre-stress, 15 minutes post-stress, 30 minutes post-stress, and 60 minutes post-stress. Yeast colonies were picked from YPD plates and grown overnight in 2 mL of sc+glu media at 30℃ until saturation. The next day, cells were back-diluted 1:200 into 25 mL of fresh, prewarmed sc+glu media in a 250-mL flask and grown for 6 hours at 30℃. Before washing the cells in preparation for ethanol stress exposure, 1.5 mL of cells were removed and centrifuged at 20,000 rcf for 1 minute to pellet. The supernatant was removed and the pellet was immediately frozen and stored at -80℃. This sample is the growth time point, used to determine gene expression levels in optimal growth conditions. The rest of the 23.5 mL of cells were then washed and stressed as described above in "Stress experiments". From the 600-microliter pre-stress cell suspension, 150 microliters were removed, pelleted, and frozen, similar to the growth time point. This sample is the pre-stress time point, used to determine gene expression levels immediately prior to stress exposure. The cells were recovered from the ethanol stress in deionized water. During this recovery period, 1.5 mL of cells were removed at 15 minutes,

30 minutes, and 60 minutes to pellet and freeze. All samples were stored at -80℃ until preparation for sequencing.

Experimental setup for fitness profiling of a pooled haploid yeast deletion library

Diagrammatically shown in Figure 2A, the pooled yeast deletion library was thawed from a frozen stock and grown in 25 mL of sc+glu media in a 250-mL flask at 30℃ until mid-log phase. Samples were then washed and stressed as described above in "Stress experiments". 150 microliters of the pre-stress cell suspension were saved as the pre-stress time point. The post-stress cell suspension was subjected to an outgrowth period in sc+glu media at 30℃ until the cells reached early log phase. Then were then saved as the post-stress time point. All experiments were performed with 24.5% ethanol and in triplicates.

Experimental setup for laboratory evolution

Diagrammatically shown Figure 3A, yeast colonies were picked from YPD plates and grown overnight in 2 mL of sc+glu media at 30℃ until saturation. The next day, cells were back-diluted 1:200 into 25 mL of fresh, prewarmed sc+glu media in a 250-mL flask at 30℃ and grown for six hours. The cells were then washed and stressed, and fraction survival was also calculated as described above in "Stress experiments". From the tenfold diluted post-stress cell resuspension, 500 microliters were added to 2 mL of sc+glu media and regrown in 30℃ until saturation, thus starting the next round of evolution. Whenever cells were back-diluted 1:200 for the six-hour growth, 300 microliters of cells were frozen and stored in -80℃ to capture the cellular state at every round of evolution. All experiments continued for 10 - 12 rounds of evolution. Three parallel lines of evolution were performed for any given strain background and stress condition. At the end

of the evolution experiment, one evolved population from each replicate line was picked for further analysis and for sequencing.

The experiment was designed in such a way as to select for both stress defense as well as fast growth. While certain laboratory evolution studies use nutrient-limited media during the recovery period after stress to limit selection for strains that grow faster or recover from the stress faster (Boer et al., 2008), we wished to avoid accumulation of mutations that cause generic slow growth, which would be a trivial solution to lethal stress survival given the established link between growth rate and stress defense. This strategy favored mutations that increased survival without compromising growth rate or recovery time. Since slow-growing mutations will be positively selected for during stress exposure and negatively selected for during recovery, there was non-monotonic and cyclical patterns of survival. As such, the last round of evolution was not necessarily when the evolved cells survive the best to lethal stress. Upon noticing this observation, we altered our method for picking the evolved population to use for sequencing and other analyses. We picked the evolved population that was highest for surviving lethal stress based on the fraction survival data from the platings instead of just using the evolved population from the last round of evolution.

Each evolved population was tested for stress survival under the same stress used to evolve them, using the protocol described above in "Stress experiments". The population was then streaked on sc+glu plates and individual colonies were picked to test for survival. We avoided picking the smallest colonies so as to not bias for slow growing cells. At least six colonies from each population were tested for survival. If at least three out of the six survived lethal stress higher than the unevolved parental population, then the three highest surviving colonies would be sent for whole-genome sequencing. Only one population fit this criterion. For all other evolved lines,

additional colonies would be tested for stress survival either until three showed higher survival than the parental strain or until about twenty colonies were tested. At this point, even if there were less than three colonies that showed higher survival, whichever colonies that did survive better was saved for sequencing.

RNA-seq of transcriptional profiling experiment

RNA was isolated using the YeaStar RNA Kit (Zymo Research). Between step 5 and 6 of the protocol, the sample was treated with DNAse I (Sigma-Aldrich). An in-column DNAse digestion was performed according to Appendix A of RNA Clean & Concentrator-5 (Zymo Research). From the isolated RNA, rRNA was removed using the Ribo-Zero rRNA Removal Kit (Illumina). Samples were barcoded and prepared for sequencing using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs). All samples were pooled and sequenced using a NextSeq 500 sequencer (Illumina).

Sequencing of fitness profiling experiment

DNA was isolated with the YeaStar Genomic DNA Kit (Zymo Research). Each gene deletion in the pooled deletion library were then amplified for sequencing. For a given gene deletion, it is replaced, immediately downstream of the start codon, with an 18-nucleotide universal priming site U1 (GATGTCCACGAGGTCTCT), a unique 20-nucleotide TAG sequence, another 18-nucleotide universal priming site U2 (CGTACGCTGCAGGTCGAC), a 1537 base pair KanMX cassette that contains the *KAN* gene, a 19-nucleotide universal priming site D2 (CGAGCTCGAATTCATCGAT), a different unique 20-nucleotide TAG sequence, a 17-nucleotide universal priming site (CTACGAGACCGACACCG), and ends with a TAA stop

codon, which replaces the normal stop codon for that gene. The *KAN* gene confers resistance to the antibiotic kanamycin in bacteria and the antibiotic geneticin in yeast. Since not all gene deletions had an annotated downstream TAG, only the upstream TAG was sequenced.

In order to use Illumina platforms for sequencing, Illumina compatible adapter were required to be added to the TAG sequence. We decided to use the Illumina Truseq adapters. The two universal priming sites (U1 and U2) were used to add the Truseq adapters. This was performed using two rounds of PCR that resulted in a final sequence of the form AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC TN[6-11]ATGGATGTCCACGAGGTCTCTNNNNNNNNNNNNNNNNNNNCGTACG CTGCAGGTCGACAGATCGGAAGAGCACACGTCTGAACTCCAGTCACNNNNNNATCT CGTATGCCGTCTTCTGCTTG. The light green sections represent the 5' and 3' ends of the Illumina Truseq adapters, with the dark green sequence indicating a specific sequence for different samples for multiplexing. We used Truseq adapters 2, 4, 5, and 6. The purple, yellow, and red sequences represent the U1, TAG, and U2 sequences, respectively. The gray section indicates custom-made internal adapters (ATCACG, TTAGGCG, ACTTGACG, GATCAGTAG, TAGCTTACAG, GGCTACGAGTG) for additional multiplexing. These internal adapters were made in such a way that as to minimize failure during sequencing due to overabundance of one type of nucleotide (Figure 2.1). All samples were pooled and sequenced using a NextSeq 500 sequencer (Illumina).

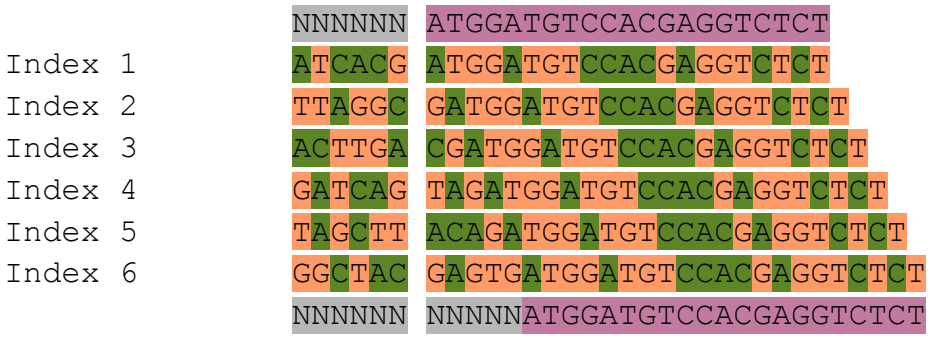


Figure 2.1 Color balance considerations for sequencing
The Illumina sequencers use two different colored lasers, one for G/T and one for A/C. At least one or two nucleotides corresponding to a given color must be read for the first several bases to prevent the run from failing. The six staggered indices (gray) shown here ensures color diversity at each position. The purple section is the U1 priming site. Figure adapted from Panos Oikonomou.

Whole-genome sequencing of laboratory evolution experiment

DNA was isolated using the YeaStar Genomic DNA Kit (Zymo Research). Isolated DNA was then barcoded and prepared for sequencing using the Nextera XT DNA Library Prep Kit (Illumina). All samples were pooled and sequenced using a NextSeq 500 sequencer (Illumina).

Pre-processing of sequencing results

All sequencing reads were first separated based on sample number using bcl2fastq conversion software (Illumina). For RNA-seq and whole-genome sequencing data, sequencing reads were clipped to remove Illumina adapter sequences (AGATCGGAAGAGC) using cutadapt (Martin, 2011). They were then trimmed using Trimmomatic 0.33 (Bolger et al., 2014) to remove end bases with a quality score below three and retain only the part of the read that has quality score above fifteen using a four-base sliding window. Reads that were less than ten base pairs as a result of this trimming were discarded. For the deletion library data, a second demultiplex step, using a custom-written script, was performed after the initial bcl2fastq to separate samples based on the internal staggered indices (Figure 2.1, gray nucleotides). By searching for the U1 and U2 primer

sequences, we were able to find the unique 20-nucleotide TAG sequence in each sequencing read. All reads that did not contain a TAG sequence were discarded.

Mapping of sequencing reads

For the RNA-seq data, reads were aligned to the reference transcriptome of *S. cerevisiae* strain S288C (https://downloads.yeastgenome.org/sequence/S288C_reference/orf_dna/) using kallisto (Bray et al., 2016). The counts for each gene from the kallisto output were expressed in the normalized form, transcripts per million (TPM). TPM counts for each sample were then combined into a matrix, and genes in which all samples had a raw read count below ten were discarded from further analysis.

For the deletion library data, reads were mapped to the reference file of all TAG sequences using bowtie2, default settings (Langmead and Salzberg, 2012). All reads that mapped to multiple TAG sequences were discarded. A total of 4,563 gene deletions were represented in our samples. The total number of TAG sequences was normalized across all samples, and all gene deletions in which the read count was less than ten was discarded.

For the whole-genome sequencing data, reads were mapped to the *S. cerevisiae* S288C genome using breseq (Deatherage and Barrick, 2014). Breseq uses bowtie2 to map reads to the reference genome. It then calls mutations using a standard Bayesian single nucleotide polymorphism caller, and each resulting score is corrected for multiple testing by multiplying by the total genome size. For populations that were sequenced, breseq was run in population mode, which identifies mutations present at intermediate frequencies in the population. Mutations completely fixed in the population would still be identified as 100% frequency (Deatherage and Barrick, 2014).

Clustering of RNA-seq data and functional category analysis

For clustering, all genes in which the coefficient of variation across all time points was below 0.2 were removed. The remaining genes were clustered in R using k-means, with cluster sizes of 10, 20 and 40. The genes and their cluster designations were used as input to iPAGE (Pathway Analysis of Gene Expression) to identify the likely pathways that are overrepresented in each of the clusters (Goodarzi et al., 2009). We decided to use a cluster size of 10 for all downstream analyses because the larger cluster sizes over-clustered the data and had clusters in which no significant pathways from iPAGE were assigned to them. To expand the list of relevant pathways, iPAGE analysis was also performed on each of the individual post-stress time points after being mathematically zero-transformed by the pre-stress time point. The full list of pathways was then filtered for only significant terms ($p<0.001$). The list of pathways was large and had many redundant terms, so it was summarized into functional categories using REVIGO (Supek et al., 2011).

Enrichment/depletion and functional category analysis

DEseq2 was used to determine gene deletions that are significantly enriched or depleted after stress survival (Love et al., 2014). The resulting enrichment scores were used as input to iPAGE. The common functional categories between the yeast deletion library data and the RNA-seq data were then compared to determine if any correlation existed between the two datasets.

Validation of evolved strains

One evolved population from the laboratory evolution experiment showed a higher survival compared to the other evolved populations. The highest surviving colony from that population was used for testing cross-resistance to other stresses and also survival at other ethanol concentrations. The mutations from that colony, as determined from the whole-genome sequencing results, were recreated in the wild-type background to test the effect of each mutation alone.

We created the mutations using a CRISPR-Cas9-mediated genome editing method described on Ryan and Cate, 2014 and Ryan et al., 2016. We first identified a protospacer adjacent motif (PAM) sequence, which is NGG, close to the site where we wanted to create the mutation. The 20 base pairs immediately upstream of the PAM sequence, called the guide sequence, was cloned into a pCAS plasmid. The original protocol, using ligation-independent cloning, created mutations in the pCAS plasmid and had low efficiency, so we used a modified version of pCAS (courtesy of the Lorraine Symington Lab) in which we used restriction cloning to insert the guide sequence into the pCAS plasmid. Digestion of the modified pCAS plasmid with the restriction enzymes ZraI and XbaI and ligation with the annealed oligomers TTT[20-nucleotide guide sequence]GTTTTAGAG and CTAGCTCTAAAAC[20-nucleotide guide sequence]AAA allowed the correct guide sequence to be successfully cloned into the pCAS plasmid. The repair sequence was designed by ensuring at least 50 base pairs of homology upstream and downstream of the mutation and PAM sequence. Since one of our mutations was not within 20 base pairs of a PAM sequence, we needed to create a synonymous mutation immediately upstream of the PAM to ensure that the CRISPR-Cas9 did not cut the DNA again. Fortunately, this mutation created an early stop codon, so the mutation created would likely not be translated. We then co-transformed our pCAS plasmid and linear repair DNA into our wild-type yeast cells. The original protocol

suggested making yeast competent cells for the transformation, but this had too low transformation efficiency. Instead, we adapted the LiAc-PEG-ssDNA method for our transformation (Gietz and Woods, 2002), using one microgram of pCAS plasmid and 2.5 micrograms of linear repair DNA.

GLOBAL RESPONSES OF YEAST CELLS TO LETHAL AND THRESHOLD LETHAL ETHANOL STRESS

SUMMARY

Yeast respond to a diverse array of mild stresses by initiating stress-specific and general gene expression programs to both survive the stress and also protect against subsequent stronger doses of stress. How yeast adapt to and survive acute exposure to lethal stresses, however, is poorly understood. Here, we use transcriptional profiling via RNA-seq and fitness profiling of the pooled yeast deletion library to discover and characterize existing and novel genes and pathways that are important in responding to and surviving ethanol stress. We find many pathways that are also significantly enriched in the environmental stress response as well as pathways known to have ethanol-specific stress responses. We also discovered novel functional categories such as chromosome condensation and spore wall assembly, known only to play a role in diploid yeast, indicating a possible adaptive role of translational repression in the ethanol stress response.

INTRODUCTION

The majority of yeast stress studies use stress levels well under the lethal threshold. This includes all the stresses used in the environmental stress response (Gasch et al., 2000; Causton et al., 2001). While some of the genes and pathways important to the response to sublethal stress may be similar to the response to lethal stress, there is likely to be novel mechanisms and pathways at play for lethal stresses. Even between two different sublethal ethanol concentrations, at 8% and 11% ethanol, there were significant differences in the genes that respond to the stress (Kubota et al., 2004; Yoshikawa et al., 2009). To study the genes and pathways important for lethal stress survival, we used two global approaches: transcriptional profiling using time-course RNA-seq and fitness profiling utilizing the pooled yeast deletion library.

Gene expression profiling using high-throughput methods has been used for decades, starting with complementary DNA microarrays (Schena et al., 1995) and greatly improved with the advent of RNA-seq (Nagalakshmi et al., 2008). RNA-seq is powerful in that it allows the entire transcriptome to be studied in many different facets, such as allele specific expression, gene fusions, non-coding transcripts, and detection of alternative splicing (Wang et al., 2009; Roy et al., 2011). Many dynamic biological processes such as the recovery from sublethal stress (Gasch et al., 2000; Causton et al., 2011) have to be captured continually in a time course experiment. The data needs to be captured at several time points during the recovery phase in order to recapitulate the entire stress response. In contrast to non-time-series RNA-seq data, time course RNA-seq data is more complex to process and requires different methods (Spies and Ciaudo, 2015). Unbiased pattern-based approaches, such as unsupervised clustering, are common methods used to analyze time courses data. Gasch et al., 2000, used hierarchical clustering for their time course data. A comprehensive study comparing four commonly used clustering algorithms using 52 gene expression datasets found that k-means/k-medoids ranked among the best methods for analyzing time-course RNA-seq data (Jaskowiak et al., 2014).

Determining the effects of a gene deletion is another fundamental approach to studying how a gene functions under exposure to stress. Using the haploid yeast deletion library, in which each non-essential gene is deleted and uniquely identified using a DNA barcode, allows the library to be pooled so that growth and fitness contributions of every non-essential gene can be studied in parallel (Giaever et al., 2002). While the early pooled yeast deletion libraries were quantitatively studied by hybridizing to high-density oligonucleotide arrays, improved methods such as barcode sequencing allowed newer high-throughput sequencing platforms to be used to quantitatively measure the fitness effect of gene deletions (Han et al., 2010).

Here, we performed transcriptional profiling via RNA-seq on haploid wild-type yeast cells at threshold lethal stress to determine the genes and pathways important for recovery from the stress. To further explore which of these functional categories may play an adaptive role in stress survival, we performed fitness profiling using the pooled yeast deletion library, which contains deletions of all non-essential genes in haploid yeast. We were able to determine which gene deletions helped or hurt survival under lethal ethanol stress.

RESULTS

Cellular response at the threshold of lethal ethanol stress

We developed an acute lethal stress paradigm in which haploid yeast cells were treated with a brief 2-minute exposure to a range of ethanol concentrations, from 19% to 26% ethanol. This brief exposure minimized the possibility of a direct transcriptional response during the period of stress. In this way, we created conditions in which the cellular state immediately prior to the stress and the longer-term transcriptional responses following the short period of stress were the dominant contributors to survival. As can be seen from the stress-survival curve in Figure 3.1, ethanol exposure above 20% causes lethality as determined by the fraction survival of CFUs. Furthermore, survival drops exponentially for concentrations above 20%, reaching 10^{-5} at 26% ethanol.

Global transcriptional response at the threshold of lethal ethanol stress

In order to gain better insights into the cellular pathways that may contribute to acute ethanol stress survival, we carried out transcriptional profiling using RNA-seq. To minimize any

confounding effects from dying cells, we chose to determine transcriptional responses at the threshold levels of lethality corresponding to 20% ethanol. The transcriptional state of haploid yeast cells (strain BY4741) were monitored starting prior to stress exposure and followed until 60 minutes after stress exposure, sampling at various time points during the recovery from stress (Figure 3.2A).

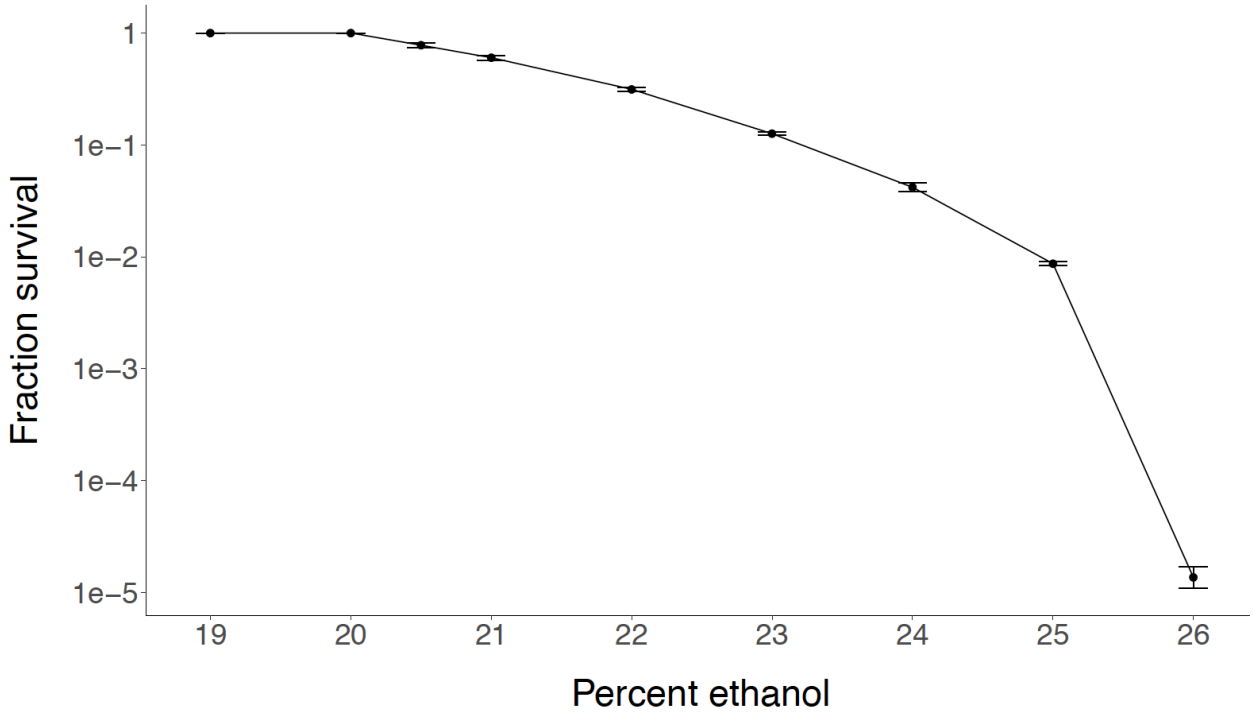


Figure 3.1 Yeast survival under a range of ethanol concentrations
The fraction survival of yeast CFUs exposed to two minutes of ethanol stress was tested at a range of ethanol concentrations, from 19% to 26%. At 20% and below, there is 100% survival. Above 20% ethanol, the fraction survival drops in a dose-dependent manner with increasing ethanol concentrations.

The samples collected from the experiment were analyzed by RNA-seq. Comparison of global transcriptional states between each post-stress time point and the pre-stress time point revealed transcriptional changes that accompany any adaptation following stress. Figure 3.2B shows the number of genes that exhibit a twofold change in expression at each post-stress time point compared to the pre-stress time point. The number of genes that showed a twofold change is

greatest at the 15-minute time points, suggesting that the peak response to ethanol stress occurs early. At the early time points (15 and 30 minutes), there are 4.5 times as many genes with a twofold decrease than a twofold increase, indicating that the transcriptional response to ethanol stress is largely driven by a global downregulation of gene expression.

To discover dominant patterns of gene expression following stress, we carried out unsupervised clustering of genes using the k-means algorithm. The gene expression patterns of the ten clusters are shown in Figure 3.2C, demonstrating both induced and repressed groups of genes. The application iPAGE (Goodarzi et al., 2009), a pathway discovery algorithm, revealed that these expression clusters were significantly enriched in various functional categories and biological processes (Figure 3.3, Supplementary Figure 1). There are four clusters (3, 4, 6, 8) that show decreased expression over the entire recovery period. Ribosome biogenesis, ribosomal proteins, and translation (in clusters 3 and 6) are three well-documented functional categories, part of the environmental stress response, that decrease in response to a variety of stresses (Gasch et al., 2000, Causton et al., 2001). The decreased expression of genes encoding for electron transport chain and oxidative phosphorylation (cluster 4), which occurs in the mitochondria, is consistent with studies that show deletion of mitochondrial genes increase survival during exposure to a variety of lethal stresses (Zakrzewska et al., 2011).

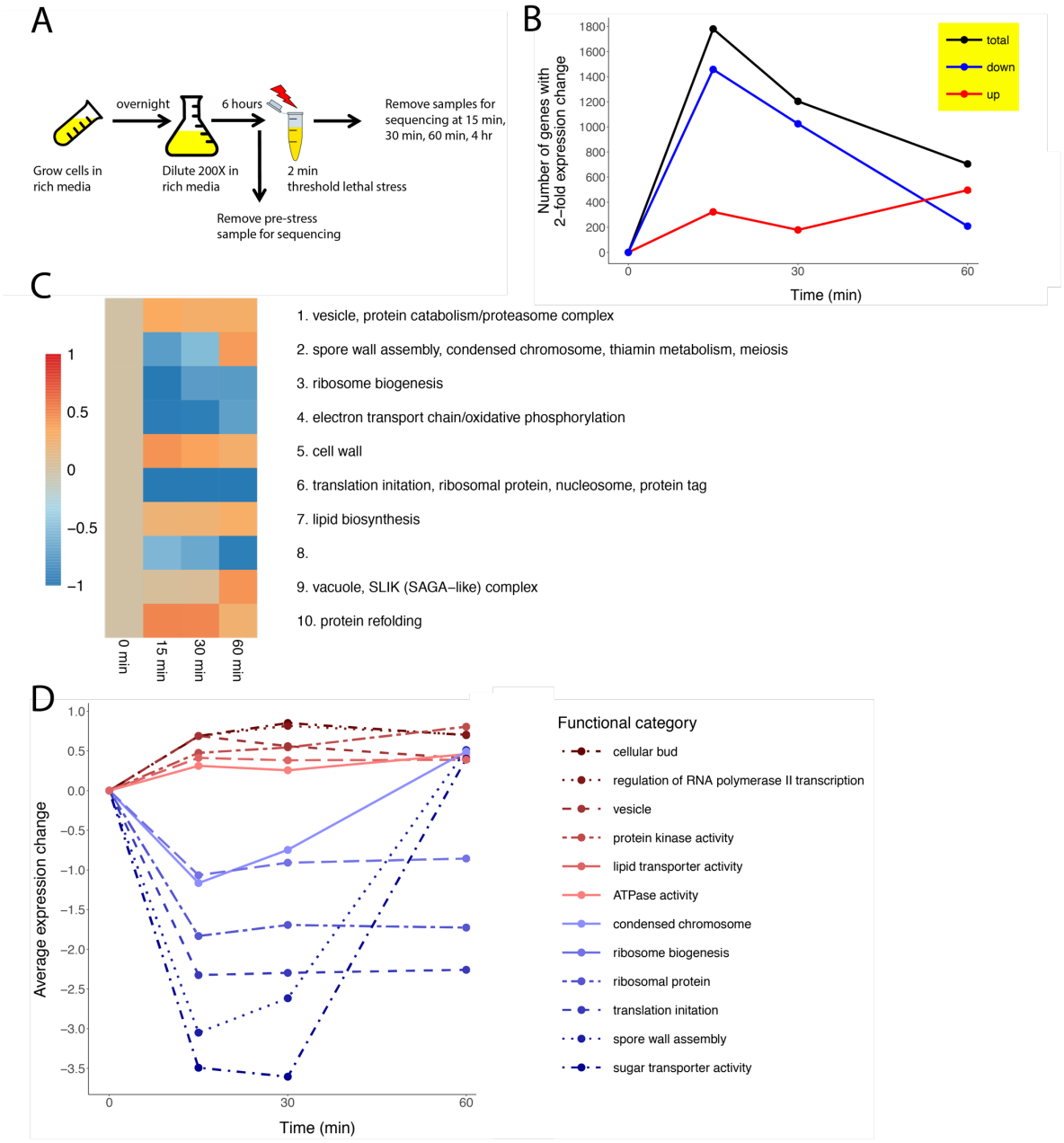


Figure 3.2 Transcriptional analysis of response to ethanol stress

(A) Procedure used to perform stress assay on yeast cells and to collect samples for RNA-seq. (B) Number of differentially expressed genes for each post-stress time point relative to the pre-stress time point, measured as the number of genes with a 2-fold expression change. The total number of differentially expressed genes was broken down into those with a 2-fold downregulation and those with a 2-fold upregulation. (C) Clustering of the time-course gene expression data using k-means with 10 clusters. Gene ontology analysis was then performed on each of the clusters, and significant functional categories ($p<0.001$) are shown for each cluster. (D) Average expression from the 12 most significant gene ontology categories in the 15- or 30-minute post-stress time point. Red colored lines indicate upregulated genes and blue colored lines indicate downregulated genes within the early post-stress time points.

Cluster 10, enriched for protein refolding, shows the strongest upregulated response during the early time periods, suggesting the possible adaptive role of protein refolding by heat shock proteins in the adaptation to ethanol stress. This is consistent with previous studies that have shown induction of heat-shock proteins in response to mild ethanol stress (4% - 10% ethanol) (Piper et al., 2004). Cluster 7 shows genes upregulated in lipid biosynthesis, more specifically, sphingolipid biosynthesis. Membrane lipid composition is known to play an important role in sublethal ethanol and heat stress (Swan and Watson, 1999), and increased sphingolipid biosynthesis has been shown to increase tolerance to heat stress in yeast (Cowart and Obeid, 2007). Other studies have also shown that vesicular and vacuolar transport are both important in yeast survival to mild ethanol stress (D'amore et al., 1989; Walker, 1998; Kubota et al., 2004; Yoshikawa et al., 2009; Charoenbhakdi et al., 2016; Navarro-Tapia et al., 2016), represented here in clusters 1 and 9, respectively. While upregulation of vesicular genes occurs shortly after exposure to ethanol, induction of vacuolar functions shows a delayed induction, peaking closer to 60 minutes after the onset of ethanol stress.

In order to capture the broadest possible set of pathways modulated during the adaptation process, we compared the gene expression at each of the three post-stress time points to the pre-stress time point and discovered gene ontology classes that are significantly enriched in them. The complete list of these categories can be found in Supplementary Table 1. The average gene expression from the 12 most significant functional categories modulated during the early time points (15 and 30 minutes) are shown in Figure 3.2D. Whereas lipid biosynthesis was significant in one of the upregulated patterns in Figure 3.2C, here lipid transporter is also shown to be significantly induced, further indicating its significance in the ethanol stress response. Regulation of RNA polymerase II transcription (Choder and Young, 1993) and protein kinase activity (Ho et

al., 2018) are both categories that have increased expression under certain stresses. On the contrary, deletion of RNA polymerase II genes has also been associated with higher survival, but this was found to mostly be a growth rate effect (Zakrzewska et al., 2011).

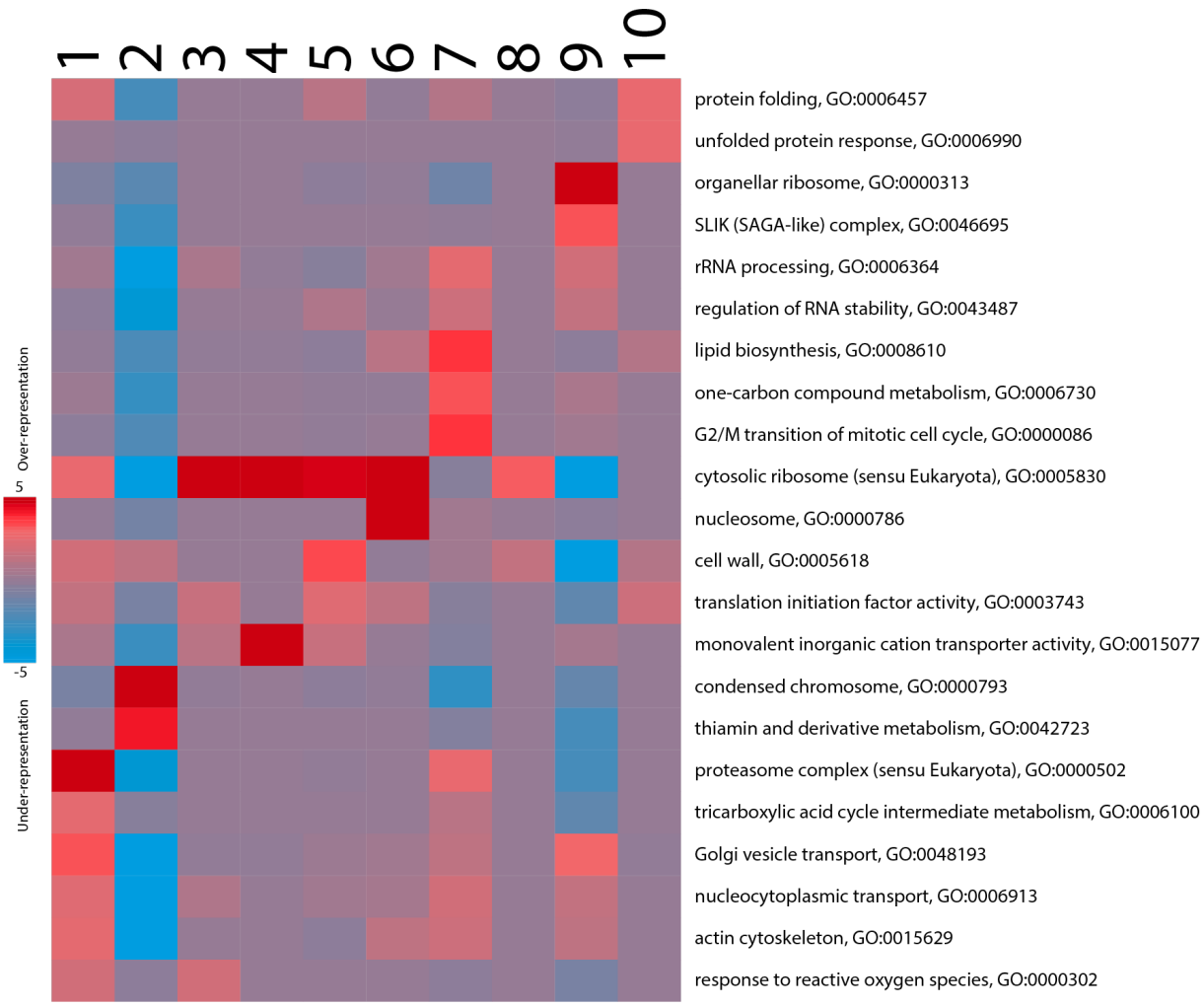


Fig 3.3 Over- and underrepresented pathways in each of the expression clusters
This is a summarized list of 22 representative pathways in each of the 10 clusters. The full list, which we used for downstream analysis, contains 202 pathways and is shown in Supplementary Figure 1. Overrepresented pathways are shown in red and underrepresented pathways are shown in blue.

The ATPase activity category consists of some genes that code for heat shock proteins, including the Hsp70 family, all of which have been shown to protect against mild ethanol stress (Piper et al., 1994). The ATPase activity category is also composed of genes that code for vacuolar H+ ATPase, required for tolerance to straight chain alcohols through vacuolar acidification

(Charoenbhakdi et al., 2016). The category that is upregulated the greatest during the early time points, cellular bud, is composed of genes involved in creating a protuberance from the mother cell to create a daughter cell (includes genes with functions in polarized growth, localization of proteins to either mother or daughter cell, cytokinesis, cell wall formation), which have not been previously implicated in ethanol stress adaptation.

The functional category that shows the greatest downregulation is sugar transport activity. Other transport processes, such as ammonium transport and organic acid transport, are also downregulated in response to ethanol stress in our data to a lesser extent. Transport processes play an important role in the ethanol stress response, with some reports of general inhibition of transport processes (Leão and Van Uden, 1984) and others reporting an increase of transport processes, specifically hexose transport (Chandler et al., 2004) under mild ethanol stress. Kubota et al., 2004 found that deletion of transport genes hurt survival when yeast cells were exposed to 11% ethanol. Vesicle-mediated transport was increased in our analysis, suggesting both an increase and decrease of transport processes at threshold lethal ethanol stress, depending on the substrate being transported.

Two categories, spore wall assembly and condensed chromosome, show an initial decrease during the early recovery period with a subsequent increase in gene expression after 60 minutes of recovery (Figure 3.2C, cluster 2). Chromosome condensation occurs during meiosis in yeast (Yang et al., 2006). Meiosis and spore formation are processes that are specific to diploid yeast (Wagstaff et al., 1982) and their modulation in response to ethanol stress in haploid yeast suggests their potential adaptive value in response to acute lethal stress.

Contribution of all non-essential genes to survival under acute exposure to lethal ethanol stress

Transcriptional responses to extreme stress can reflect pathways that may be adaptive. Alternatively, these gene expression dynamics may reflect non-adaptive or even maladaptive responses that may only be of values within the organism's native habitat (Tagkopoulos et al., 2008). In order to systematically determine the contribution of yeast genes to acute lethal stress, we carried out fitness profiling of a pooled haploid yeast deletion library upon exposure to lethal ethanol stress (Figure 3.4). We use a lethal level of ethanol, specifically at 24.5% ethanol, the concentration at which 1% of the wild-type population survives after a two-minutes exposure. The inclusion of a post-stress outgrowth phase reduced the likeliness that some mutants may achieve better survival due to a severe growth defect. This haploid library contains about 4,800 deletions of most non-essential genes, with each gene deletion represented by a unique 20-nucleotide barcode sequence. Comparing the abundance of post-stress gene deletions to pre-stress gene deletions, we were able to systematically quantify the effects of each gene deletion on survival.

This analysis revealed gene deletions that significantly increased or decreased survival. The list of the top gene deletions with positive and negative fitness effects are shown in Table 3.1. For the gene deletions that help survival, five out of the top 20 genes are ribosomal subunits, three are mitochondrial proteins, and two are in the TOR pathway. The increased survival seen from deletion of ribosomal genes and mitochondrial genes are consistent with prior yeast deletion studies showing the same effect to other stresses (Zakrzewska et al., 2011; Welch et al., 2013). Decreased expression of TOR pathway components is well known to increase survival universally under stress conditions (Powers et al., 2006; Wei et al., 2009; Pan et al., 2011). If the list is expanded to the top 100 gene deletions, many more ribosomal genes show up. For the gene

deletions that hurt survival, six out of the top 20 genes have vacuolar functions. This suggests that vacuolar functions are essential for yeast survival under lethal ethanol stress.

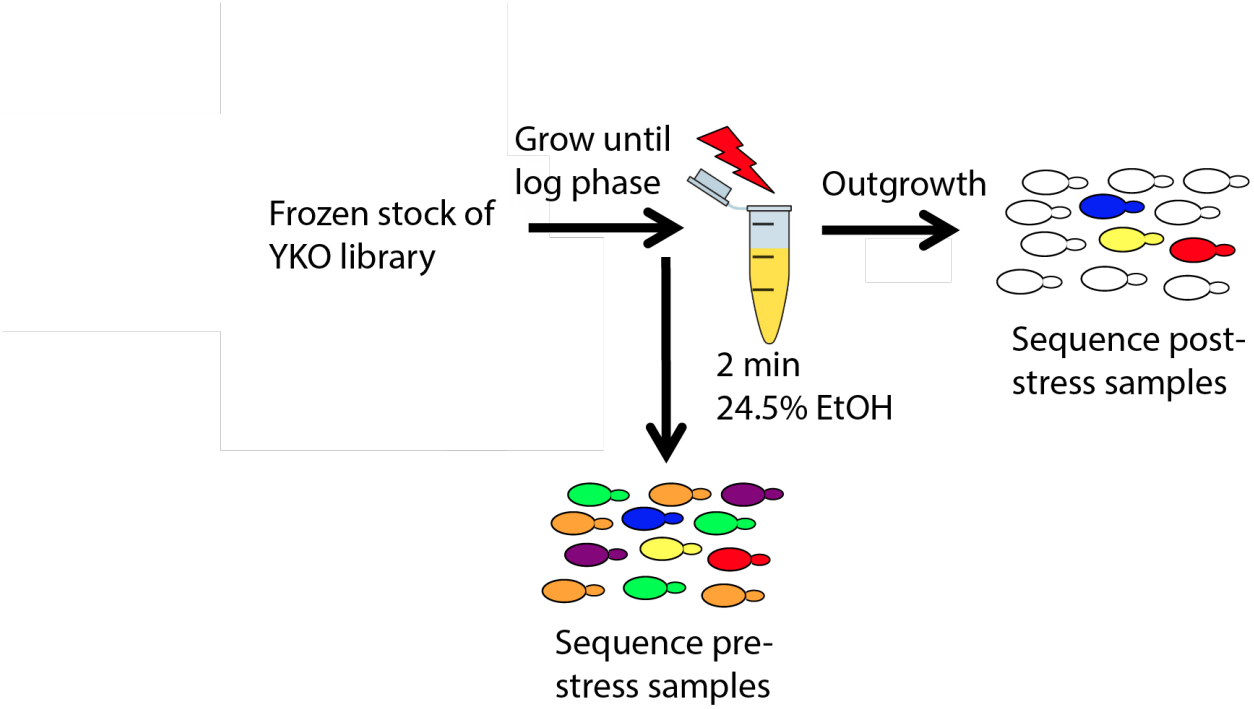


Figure 3.4 Procedure used to perform stress assay on yeast cells and to collect cells for barcode sequencing

We next performed iPAGE analysis on the full range of fitness scores in order to determine whether specific pathways were enriched among gene deletions that changed fitness (Figure 3.5). We found significant enrichment for ribosome and translation-related genes at both the positive and negative ends of the fitness effect distribution. In particular, there were many non-essential ribosome and translation-related genes whose deletion substantially improved survival. The fitness benefit of deleting non-essential translation-related functions is consistent with the strong transcriptional repression of translation immediately following lethal stress (Figure 3.6). This suggests that repression of translation is indeed an adaptive response to lethal stress. In addition to translational repression, we wanted to explore whether there was a correlation between the maximal transcriptional change of a gene across the post-stress time course and the fitness effect

of a deletion in that gene. Plotting all significant positive and negative fitness scores of gene deletions again their maximal post-stress expression fold change, we noticed that about 85% of all genes exhibited this anticorrelation (Figure 3.7). Many of these genes may play an adaptive role in response to lethal ethanol stress.

ENRICHED			
Gene Name	**Systematic Name**	**Score**	**Description**
AAH1	YNL141W	8.26	adenine deaminase
LTV1	YKL143W	8.05	EGO/GSE complex subunit, upstream of TOR complex
FTR1	YER145C	7.72	iron permease
RPL13A	YDL082W	7.72	ribosomal 60S subunit
SNT1	YCR033W	7.71	deacetylase, positive regulation of stress-activated MAPK cascade
AFG3	YER017C	7.59	mitochondrial metallopeptidase
RPS16A	YMR143W	7.47	ribosomal 40S subunit
TRM13	YOL125W	7.38	methyltransferase
EAP1	YKL204W	7.15	translation initiation factor, TOR pathway
RPL14A	YKL006W	6.90	ribosomal 60S subunit
RPL16B	YNL069C	6.80	ribosomal 60S subunit
PIL1	YGR086C	6.51	eisosome assembly
KSS1	YGR040W	6.11	MAP kinase

PET117	YER058W	5.86	electron transport chain
ASF1	YJL115W	4.20	nucleosome assembly factor, stress response
MRP7	YNL005C	4.02	mitochondrial ribosomal large subunit
VPS24	YKL041W	3.81	vacuolar protein sorting
CTR1	YPR124W	3.74	copper transporter
REG1	YDR028C	3.49	protein phosphatase regulator
GAT2	YMR136W	3.34	transcription factor

DEPLETED			
Gene Name	**Systematic Name**	**Score**	**Description**
VPS69	YPR087W	-12.40	vacuolar protein sorting
SIW14	YNL032W	-6.92	tyrosine phosphatase
NGG1	YDR176W	-6.81	acetyltransferase
YJL047C-A	YJL047C-A	-6.79	unknown function
DGA1	YOR245C	-6.38	acyltransferase, triglyceride biosynthesis
INO2	YDR123C	-6.34	transcription factor, phospholipid biosynthesis
IRC23	YOR044W	-6.16	unknown function
VPS13	YLL040C	-5.97	vacuolar protein sorting

SUE1	YPR151C	-5.96	degradation of cytochrome c
UGA4	YDL210W	-5.79	GABA transport, located on vacuole membrane
MON1	YGL124C	-5.70	guanine nucleotide exchange factor, located on vacuole membrane
DUR3	YHL016C	-5.64	putrescine, spermidine, urea transporter
ISN1	YOR155C	-5.53	nucleotidase, breaks IMP to inosine
AHK1	YDL073W	-5.52	scaffold protein, important in osmotic stress
YCF1	YDR135C	-5.32	glutathione transporter, located on vacuole membrane
PMC1	YGL006W	-5.28	ATPase, located on vacuole membrane
YGL015C	YGL015C	-5.25	unknown function
YGL140C	YGL140C	-5.24	unknown function
ICL1	YER065C	-5.24	isocitrate lyase, induced by growth on ethanol
TUF1	YOR187W	-5.24	mitochondrial translation elongation factor

Table 3.1 Top 20 most enriched and depleted gene deletions upon exposure to lethal ethanol stress

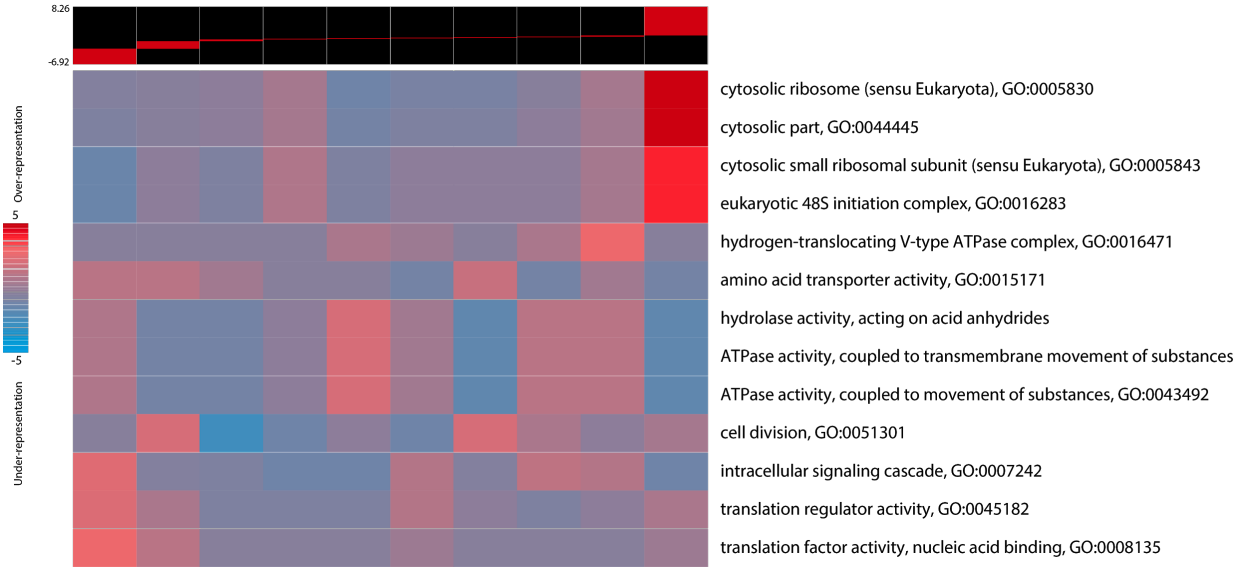


Figure 3.5 Over- and underrepresented pathways within the full range of fitness scores
The pathways that are over- or underrepresented among the fitness scores are shown. Overrepresented pathways are shown in red and underrepresented pathways are shown in blue.

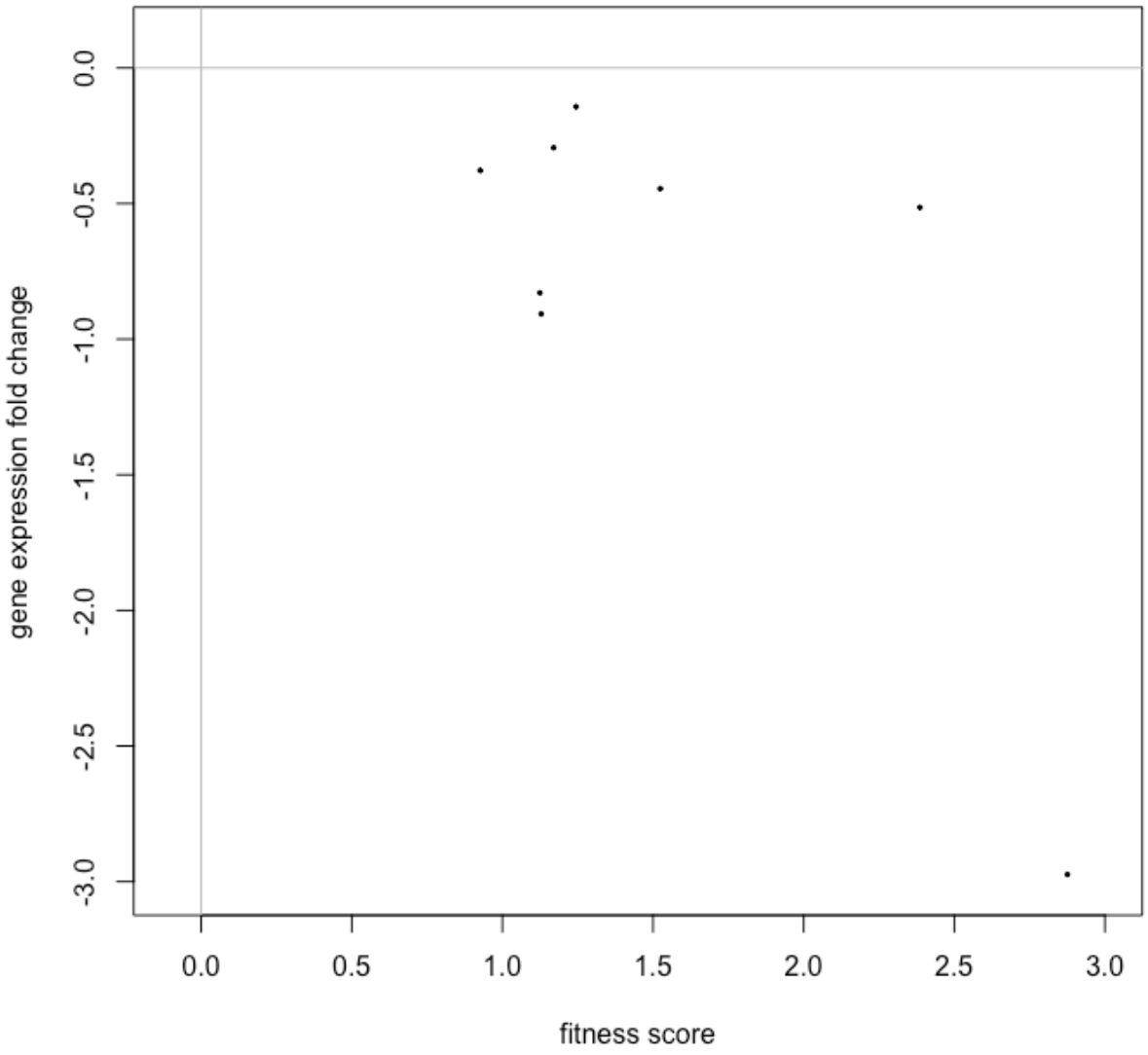


Figure 3.6 Gene expression fold change versus fitness score for translation initiation genes
The significantly enriched translational initiation genes from Figure 3.5 (rightmost column, fourth row) were used to plot their maximal post-stress expression fold change from the RNA-seq data against their fitness score from the deletion library data.

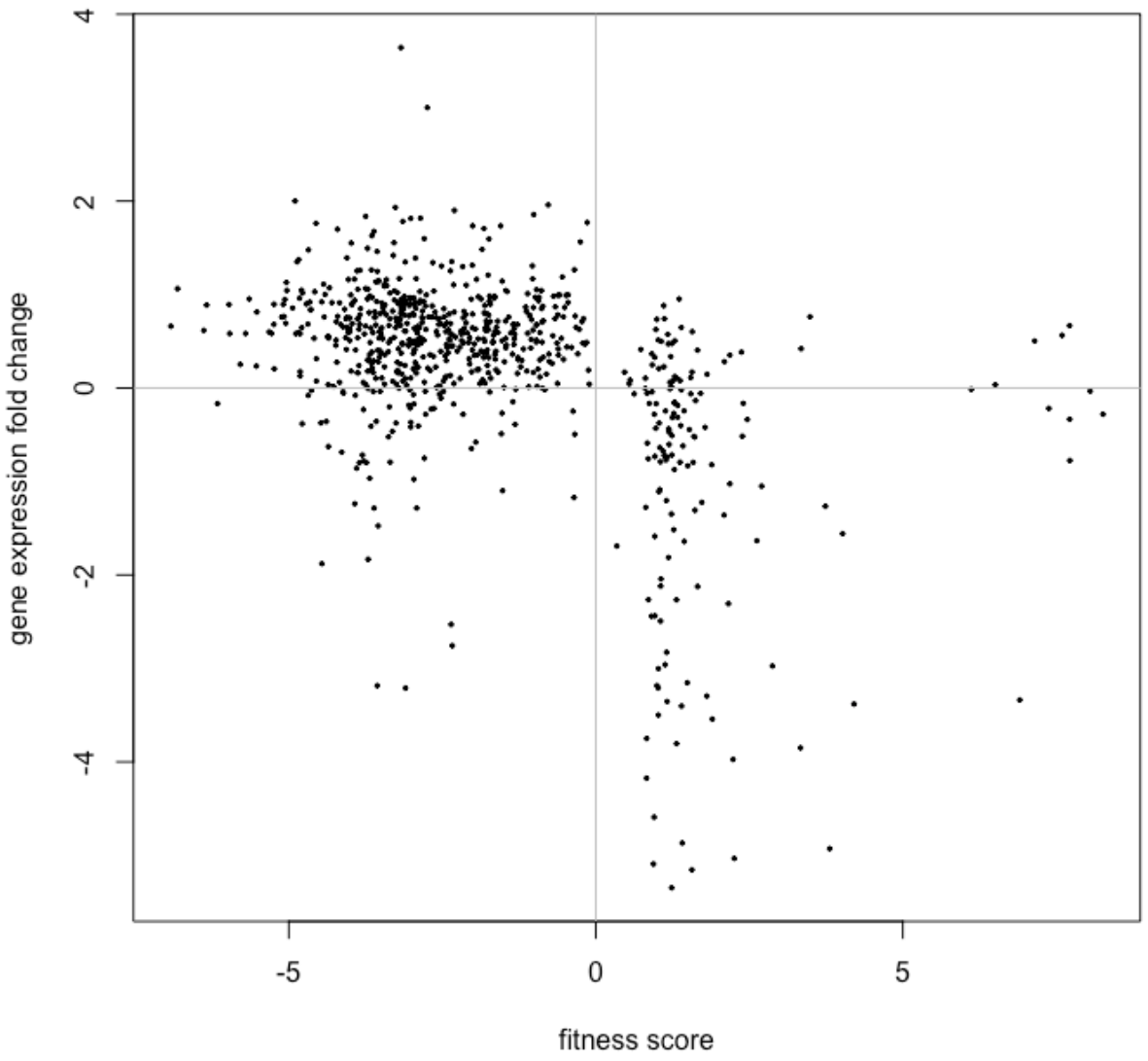


Figure 3.7 Gene expression fold change versus all significant fitness scores
All significant fitness scores ($p<0.05$, Bonferroni corrected) were plotted against their maximal post-stress expression fold change from the RNA-seq data. 85% of all points were contained within the top left and lower right quadrants.

DISCUSSION

Using two global approaches, we were able to discover many important genes and pathways important for adaptation to threshold lethal ethanol stress and for survival to lethal ethanol stress. Our method of measuring gene expression at the threshold of lethality under such a short time period was a novel approach to studying how genes respond to stress and how they allow yeast cells to recover. Nevertheless, we were able to discover pathways that were previously known to play a role in stress survival, both under general stress and under ethanol-specific responses. We found the functional categories of ribosome biogenesis, ribosomal proteins, and translation, of all which are known to be repressed in the ESR, to also be downregulated in our experiment as well, highlighting the importance of these general stress response mechanisms under ethanol stress.

Cellular membranes are targets of ethanol stress, and we found lipid biosynthesis and lipid transporter as important upregulated functional categories in the post-stress response. It has been shown that the main targets of ethanol in yeast are the plasma membrane, mitochondrial membranes, nuclear membrane, vacuolar membrane, and endoplasmic reticulum (D' amore et al., 1989; Walker, 1998). The particular genes that make up the lipid biosynthesis category in our dataset mostly include genes that code for components of the plasma and mitochondrial membranes, but there were also genes that code for components of endoplasmic reticulum membranes, vacuolar membranes, and cell wall components. The lipid transporter category in our dataset mostly includes genes that are involved in the transport of phospholipids in general. More specifically, there were a few genes that involved the transport of phosphatidylinositol. Taken together, these results underscore the importance of protecting against damage to cellular membranes due to ethanol.

In addition to pathways shared between both sublethal and threshold lethal stress, we found novel functional categories that are significantly enriched in response to ethanol stress. The categories of chromosome condensation, which occurs during meiosis, and spore wall assembly both only occur in diploid yeast. The fact that they are significantly downregulated in response to ethanol stress suggests a possible adaptive role under acute lethal stress. In diploid yeast, the stress of starvation propels yeast cells to undergo meiosis and sporulation (Nasmyth, 1982). Here, our yeast cells have undergone ethanol instead of starvation stress, but even if we assume they were in a starvation state, the decrease in meiosis and spore wall assembly genes is the opposite of what we would expect to occur in diploid yeast cells. Chu et al., 1998 performed time-course transcriptional profiling during meiosis in yeast to determine all the genes that were significantly induced during this process. They broke down the meiotic process into early, early middle, middle, and late middle stages. Out of the 69 genes in our dataset that fit the meiosis category, 13 of them were also found in the dataset by Chu et al., 1998. Strikingly, all 13 of our genes were found to be induced in the early stage of meiosis in their dataset.

Most of the genes known to be involved in the early stage of meiosis function in meiotic prophase, which consists of pairing of homologous chromosomes and recombination. Many of these genes contain a conserved site, URS1 (5'-GGCGGC-3'), in their upstream region. This sequence is recognized by Ume6/Ime1, a major transcriptional regulator of this group of early meiotic genes (Rubin-Bejerano et al., 1996; Kupiec et al., 1997). Analysis of our data using FIRE also discovered the URS1 consensus site in the genes from the meiosis category. The role of URS1 in mediating stress in yeast has been largely unstudied. It is known that in diploid yeast, URS1-bound factors modulate *HSF1* through *HSP82* during early meiotic induction (Szent-Gyorgyi, 1995). Under heat shock, protein binding at URS1 is lost (Erkine et al., 1999). This allows the

URS1-containing *HSP82* to be de-repressed and can subsequently activate *HSF1* (Uffenbeck and Krebs, 2006).

Genes that contain URS1 can be activated by *IME1*. In our dataset, it is shown that *IME1* expression mimics the pattern shown by spore wall assembly genes seen in Figure 3.2. Its role as a master regulator of meiosis is primarily as an early inducer. While this is well known to occur in diploid yeast, its role in haploid yeast is poorly understood. We do not know why its expression decreases in response to ethanol stress, but its strong downregulation suggests an adaptive role in the response to ethanol stress.

Fitness profiling revealed many gene deletions that have similar functions at both the extremes of significantly enriched and significantly depleted. While it may seem to be a waste of cellular resources to have the same functions both induced and repressed, this could indicate a finer control of that pathway under ethanol stress exposure. Specifically, as some studies have suggested, this could allow the cell to shift control from transcriptional machinery to translational machinery (Ho et al., 2018). Combining both fitness profiling and transcriptional profiling results allowed us to analyze which genes and pathways possibly played an adaptive role in response to acute lethal ethanol stress. Translational repression seems to be one such category that fits this potentially adaptive role.

It is important to note that the pooled yeast deletion library contains only non-essential gene deletions. There is the possibly that some essential genes play an adaptive role in the ethanol stress response, but we would not be able to test their deletions using this deletion library. A possible way to test those genes would be to use temperature-sensitive mutants or use a heterozygous diploid library. Hillenmeyer et al., 2008 used yeast heterozygous and homozygous deletion libraries in various chemical and environmental stress conditions, noting that 97% of all

genes are optimal for growth in some condition. Therefore, it is very likely that some essential genes would have their deletion play an important role in the ethanol stress response.

EXPERIMENTAL EVOLUTION TO INCREASE SURVIVAL TO LETHAL ETHANOL STRESS

SUMMARY

Mutational studies are important in determining genes important in the ethanol stress response. In the previous chapter, we used a deletion library to determine how nonessential loss-of-function mutations contribute to survival under lethal ethanol stress. Here, we expand the landscape of possible mutations by using experimental laboratory evolution to try and improve yeast stress survival under lethal ethanol stress. We were able to evolve a strain of yeast that displayed greater than 30% survival compared to a 1% survival of the parental wild-type strain. The pre-stress gene expression state of this evolved strain was similar to the post-stress expression state of the parental state. By being in a stress-tolerant state, the evolved strain was able to survive better under heat and hydrogen peroxide stress as well. This was achieved without a significant decrease in the bulk growth rate of the evolved strain.

INTRODUCTION

Experimental laboratory evolution is a method commonly used to gain insights into adaptive changes that may accumulate in populations under certain growth or environmental conditions. Many such experiments have been performed in *E. coli* and *S. cerevisiae* in the past few decades (Paquin and Adams, 1983a-b; Bennett et al., 1990). During the laboratory evolution process, the microorganism is grown in clearly defined conditions for time periods ranging from weeks to years, allowing for the selection of phenotypes that improve its fitness in that condition (Dragosits and Mattanovich, 2013). The longest running laboratory evolution study was conducted by Richard Lenski, having evolved *E. coli* for much longer than 50,000 generations (Sniegowski et al., 1997; Lenski et al., 1998; Cooper and Lenski, 2000). Many biological phenomena can be explored with these experiments, giving insights into the genetic basis of increased fitness (Barrick

et al., 2009), historical contingency (Blount et al., 2008), clonal interference (Kao and Sherlock, 2008), and evolutionary bet hedging (Beaumont et al., 2009).

In *E. coli* and *S. cerevisiae* evolution experiments, 100 to 500 generations usually correspond to a fitness increase of 50-100%, and in certain cases, even a 1000% increase can be achieved (Tremblay et al., 2010). It is hard, however, to predict the time scale at which a fitness increase of the population will occur. While some environments can lead to fast phenotypic increases, others, such as adaptation to pH, have low adaptive potential (Hughes et al., 2007; Dragosits et al., 2014). The fitness increase as a function of the number of selection cycles is also not linear. It seems to increase quickly within the first 100 to 500 generations below slowing down drastically (Barrick et al., 2009).

Here, we used experimental laboratory evolution to evolve a wild-type strain of haploid yeast over the course of 10 to 12 rounds of evolution. We calibrated the initial level of ethanol stress to allow only 1% of the population to survive. The surviving cells would be grown to saturation in rich media before starting the next round of evolution. We were able to evolve several parallel lines to have higher survival than the parental strain, with the best evolved strain showing greater than 30% survival under acute exposure to lethal ethanol stress.

RESULTS

The genetic diversity of the yeast deletion library enabled us to determine the extent to which loss-of-function mutations substantially improves survival of yeast cells to lethal ethanol stress. We were curious to what extent random mutations and selection may drive improved lethal stress survival and whether and to what extent the transcriptional responses of these cells may shift as a function of improved survival. To this end, we developed a laboratory experimental evolution

paradigm in which wild-type haploid yeast cells were exposed to lethal ethanol stress and then recovered in an outgrowth phase over multiple cycles of selection. In each round, cells were first grown to saturation in rich media, diluted and grown until mid-log phase, then exposed to two minutes of lethal ethanol stress. The surviving cells were then regrown to saturation, thus starting the next round of selection (Figure 4.1A). Whereas some studies use nutrient-limited media during the regrowth phase so as to minimize growth-based competition (Boer et al., 2008), we wished to avoid accumulation of mutations that cause generic slow growth which would be a trivial solution to lethal stress survival. The strategy we used favored mutations that increased survival without compromising growth rate. As such, we decided to recover the cells in rich media following lethal stress. Since slow-growing beneficial mutations will be positively selected for during stress exposure and negatively selected for during recovery, it may explain the non-monotonic increase of survival throughout the laboratory evolution process (Figure 4.2).

Experimental evolution was carried out in triplicate populations. The lethal selection was calibrated to have a baseline survival of 1%. As can be seen in Figure 4.2, we saw a substantial increase in survival frequency in multiple lineages. The best performing evolved population showed about an average 35% survival compared to the parental strain (Figure 4.1B, compare red to black). Individual colonies were then selected from this evolved population and also tested for survival. All colonies tested exhibited an average survival of at least 10% (Figure 4.1B, green). The colony with the highest average percent survival, which we will refer to as strain JY304, was chosen for subsequent phenotypic and molecular analyses. We first determined the survival advantage of strain JY304 across a range of ethanol concentrations. We saw a dramatic increase in survival across the full range of ethanol concentrations. At the highest concentration tested (26%), we saw nearly a hundred-fold increase in survival (Figure 4.1C). Remarkably, this survival

advantage was not accompanied by any significant reduction in the bulk growth rate of the population (Figure 4.1F).

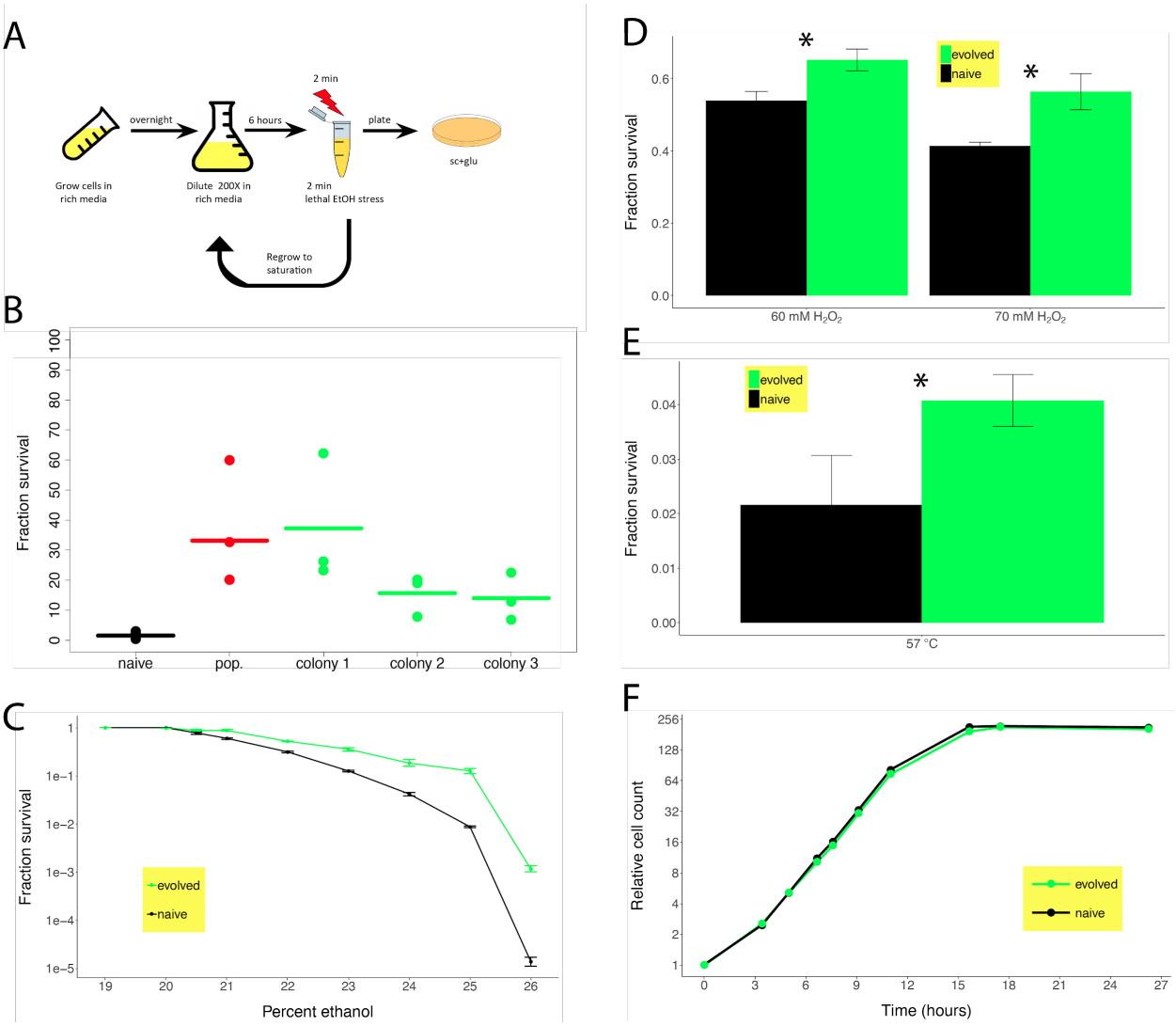


Figure 4.1 Enrichment of mutations that confer survival benefit upon acute lethal stress exposure

(A) Laboratory evolution protocol. At every round of evolution, survival was measured by plating, and experiments were stopped when survival seemed to be noticeably above the baseline survival of 1%. (B) Fraction survival of the naive wild-type strain (black), evolved population (red), and 3 distinct colonies from the evolved population (green). Horizontal lines indicate the average survival of each of the replicates. (C) Analogous to Figure S1, comparing the fraction survival of yeast at a range of ethanol concentrations, from 19% to 26%. (D) Testing cross-protection of cells evolved under ethanol stress, stressed at 60mM or 70mM hydrogen peroxide. (E) Same as D except with heat stress. (F) Growth curves of the evolved and wild-type strains.

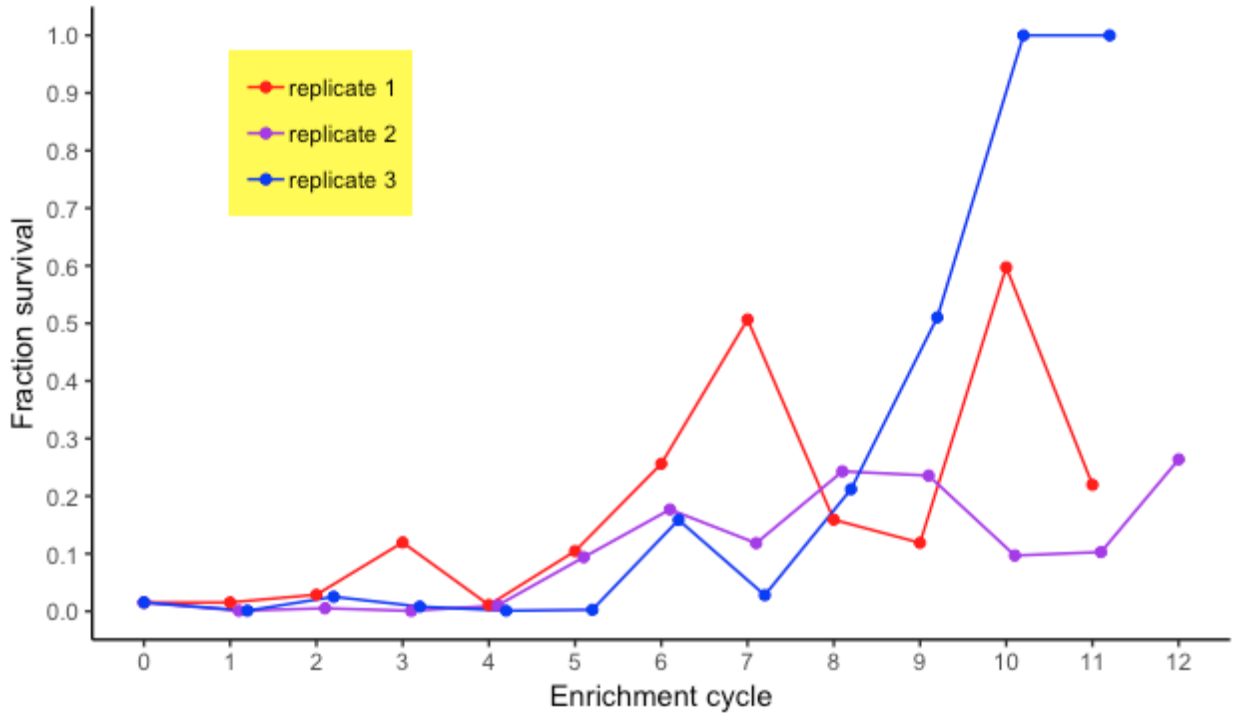


Figure 4.2 Tracking survival at each round of evolution
The fraction survival was calculated at each round of evolution for all three replicate lines of the wild-type background.

Cross-resistance of strain JY304 across orthogonal lethal stresses

The survival advantage of strain JY304 may be due exclusively to ethanol-specific advantages conferred by the underlying mutations. Alternatively, at least some of the survival advantage maybe due to more general survival effects beyond lethal ethanol stress. To see any evidence for such cross-resistance, we determined haploid yeast survival to orthogonal stresses: hydrogen peroxide (70 mM) and heat shock (57°C) for two minutes. As can be seen in Figure 4.1D-E, strain JY304 showed significantly higher survival for both stresses, suggesting that at least part of the survival advantage is generic.

Evolution of transcriptional responses in the evolved hyper-surviving strain

The evolved strain JY304 may achieve superior ethanol survival due to an improved ability to reprogram gene expression during the minutes and hours following exposure to lethal stress. Alternatively, some of the survival advantage may be due to establishment of a pre-stress cellular state that is more resistant to lethal stress. In order to determine the extent to which these differing strategies may contribute to survival, we carried out transcriptional profiling of the strain JY304 exactly as performed for the parental wild-type strain. The parental strain showed a dramatic shift in gene expression in the 15-minute post-stress period, repressing and inducing a large number of genes that led to a substantial divergence in global transcriptional state, followed by a gradual recovery of gene expression back to pre-stress patterns after an hour (Figure 3.1B). The evolved strain, however, exhibited a less dramatic shift in gene expression that was largely maintained in the post-stress period without significant recovery (Figure 4.3A).

To discover significant patterns of gene expression, we performed unsupervised clustering using the k-means algorithm with k=10. This revealed gene expression clusters, six of which showed the same enrichment in functional categories as observed in the parental strain. Three of these were upregulated (vesicle, proteasome complex, protein refolding) and three were downregulated (ribosomal protein, ribosome biogenesis, nucleosome). Although the directionality of these gene expression dynamics was similar to the parental strain, the magnitude of change was significantly different. All six categories showed a lower magnitude of change in strain JY304 compared to the wild-type strain. The overall lower magnitude of change is also evident in the comparison of heat maps (compare Figure 4.3B to Figure 3.2C).

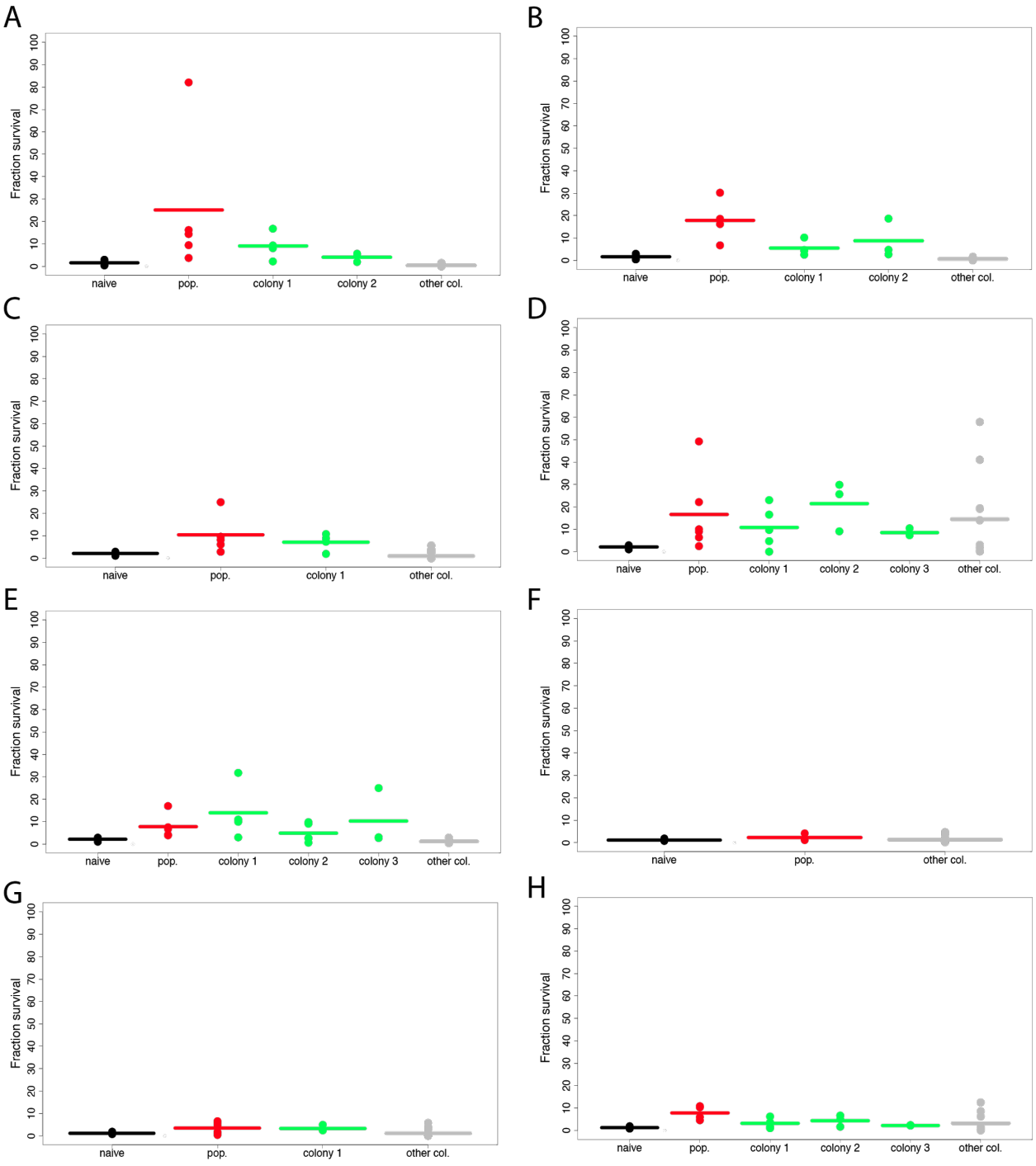


Figure 4.3 Increase in survival frequency in multiple lineages
Analogous to Figure 3B, the 8 categories are (A) wild-type replicate line 2, (B) wild-type replicate line 3, (C) ΔTCO89 replicate line 1, (D) ΔTCO89 replicate line 2, (E) ΔTCO89 replicate line 3, (F) ΔTOR1 replicate line 1, (G) ΔTOR1 replicate line 2, (H) ΔTOR1 replicate line 3. The last column in each graph, named "other col", are additional colonies tested that did not end up having a significant survival increase upon repeated testing.

We compared the global expression of genes in strain JY304 to the parental strain during exponential growth phase in rich media to determine if the pre-stress cellular state of strain JY304 is more resistant to lethal stress. Remarkably, we saw significant differences in pre-stress gene expression states with 266 genes showing greater than twofold expression change and 52 showing genes lower than twofold expression change. Table 4.1 contains a list of the 20 most highly differentially expressed genes in the both the positive and negative directions. It is worth noting that many of the most highly differentially expressed genes are of unknown function (8 out of 20 for genes with increased expression and 5 out of 20 for genes with decreased expression), including the top two genes with decreased expression. This highlights the importance of novel genes and pathways in the ethanol stress response that have yet to be elucidated. There are also categories that have genes in both the increased and decreased sets. Transposable element (four in increased, four in decreased), heat shock protein (one in increased, one in decreased), helicase (one in increased, one in decreased), and mitochondrial protein (five in increased, two in decreased) are all such examples. Having genes with similar functions being both increased and decreased in strain JY304 compared to the parental strain suggests a finer control of these specific pathways by shifting the regulation to be controlled translationally (Gasch et al., 2000). Ho et al., 2018 showed that protein production is controlled transcriptionally under normal conditions, but stress shifts it under translational control.

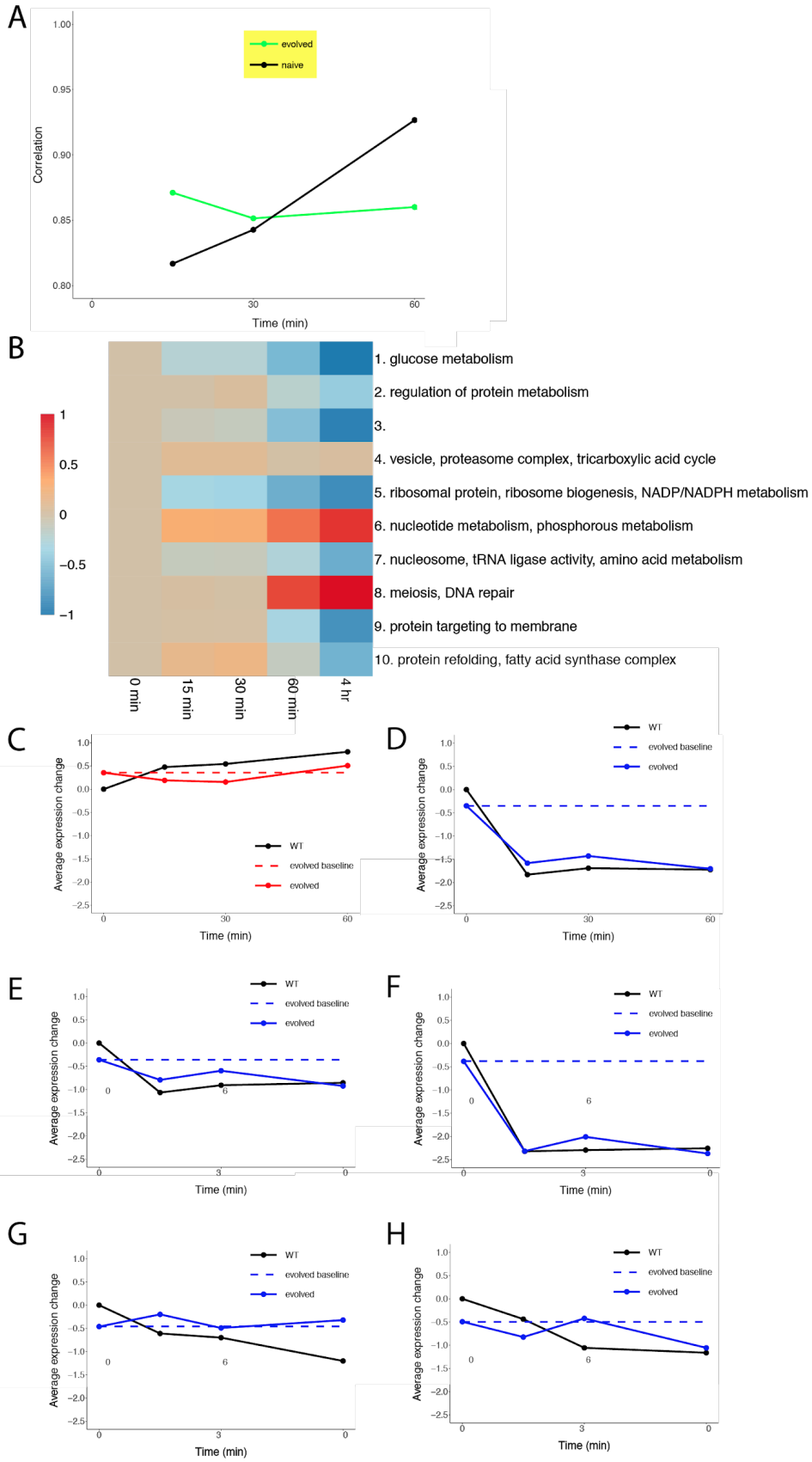

A
evolved
naive
Correlation
Time (min)
B
1. glucose metabolism
2. regulation of protein metabolism
3.
4. vesicle, proteasome complex, tricarboxylic acid cycle
5. ribosomal protein, ribosome biogenesis, NADP/NADPH metabolism
6. nucleotide metabolism, phosphorous metabolism
7. nucleosome, tRNA ligase activity, amino acid metabolism
8. meiosis, DNA repair
9. protein targeting to membrane
10. protein refolding, fatty acid synthase complex
0 min
15 min
30 min
60 min
4 hr
C
D
E
F
G
H
WT
evolved baseline
evolved
Average expression change
Time (min)

Figure 4.4 Transcriptional analysis of an ethanol-tolerant strain's response to ethanol stress

(A) Correlation between each of the post-stress time points to the pre-stress time point for both the evolved JY304 and wild-type strains. (B) Analogous to Figure 1C except on strain JY304. (C-H) Average expression of the 6 functional categories that are significant ($p<0.001$) in both strain JY304 and wild-type data and have the same directionality. The wild-type yeast data always starts at 0 at time point 0, but strain JY304 data starts at its relative expression compared to the wild-type strain during the growth phase. The 6 categories are kinase, ribosomal protein, ribosome biogenesis, translation, nucleosome, and protein tag, respectively.

INCREASED EXPRESSION			
Gene Name	**Systematic Name**	**Fold increase**	**Description**
YMR046C	YMR046C	26.7	transposable element
YKL131W	YKL131W	6.7	unknown function
YLR410W-A	YLR410W-A	6.0	transposable element
HSP32	YPL280W	5.5	heat shock protein
YBL111C	YBL111C	4.9	helicase-like protein
YHL037C	YHL037C	4.8	unknown function
YHL005C	YHL005C	4.8	unknown function
YGL235W	YGL235W	4.4	unknown function
ATP6	Q0085	4.3	mitochondrial protein
AI1	Q0050	4.2	mitochondrial protein
YPL060C-A	YPL060C-A	3.8	transposable element
SDP1	YIL113W	3.8	stress-inducible MAP kinase phosphatase, shifts location upon heat stress

YDR261C-C	YDR261C-C	3.7	transposable element
YPR136C	YPR136C	3.7	unknown function
YBR224W	YBR224W	3.6	unknown function
RRT16	YNL105W	3.5	unknown function
COX3	Q0275	3.4	mitochondrial protein
AI2	Q0055	3.4	mitochondrial protein
YPR197C	YPR197C	3.3	unknown function
COB	Q0105	3.2	mitochondrial protein

DECREASE EXPRESSION			
Gene Name	**Systematic Name**	**Fold decrease**	**Description**
YGL088W	YGL088W	75.9	unknown function
YDR545C-A	YDR545C-A	38.3	unknown function
SNA3	YJL151C	9.8	vesicular sorting of proteins to vacuole
YIL177C	YIL177C	7.9	helicase
SRB6	YBR253W	6.7	RNA polymerase II subunit
CUE4	YML101C	5.2	unknown function
YGR027W-A	YGR027W-A	4.9	transposable element

TAR1	YLR154W-C	4.7	mitochondrial protein
YLR156W	YLR156W	3.9	unknown function
COX13	YGL191W	3.8	mitochondrial protein
PGA2	YNL149C	3.6	protein transport
HSP33	YOR391C	3.4	heat shock protein
HHT1	YBR010W	2.9	histone H3, rRNA transcription
YBR064W	YBR064W	2.8	unknown function
COF1	YLL050C	2.8	golgi to plasma membrane transport
YAR010C	YAR010C	2.8	transposable element
YJR028W	YJR028W	2.8	transposable element
YPR137C-A	YPR137C-A	2.8	transposable element
RPB9	YGL070C	2.7	RNA polymerase II subunit
HTB2	YBL002W	2.5	Histone H2B

Table 4.1 Top 20 most differentially expressed genes in both the positive and negative directions in strain JY304 compared to the parental strain

Pathway analysis on the full range of pre-stress fold changes by iPAGE revealed 125 gene ontology terms with significant differences in expression ($p<0.001$), clearly showing a significant difference in the pre-stress cellular state between these two strains. The full list of 125 terms can be found in Supplementary Table 2. Many of those gene ontology terms were combined into six representative functional categories to see the dynamics in the average behavior of these gene sets

across the post-stress period (Figure 4.3, C-H). Examination of these pathway dynamics shows that in strain JY304, there was a reprogramming of gene expression that shifted its pre-stress state towards that of the wild-type at 15 minutes post-stress. For each of the functional categories that is upregulated post-stress in the parental strain, strain JY304 already has a higher expression during the growth phase (red dotted line). Likewise, for each of the functional categories that is downregulated post-stress in the parental strain, strain JY304 already has lower expression during the growth phase (blue dotted line). This suggests that the evolved strain is already in a more stress-resistant state during normal growth and is better prepared to survive an acute transition to lethal stress. Four of the six categories are known to be part of a general stress response (protein kinase, ribosomal protein, ribosome biogenesis, and translation), explaining why strain JY304 also better survives hydrogen peroxide and heat stress (Figure 4.1 D-E).

Similar to Figure 3.7, we plotted the ratio of the maximal post-stress transcriptional response of the wild-type strain to the pre-stress expression level of strain JY304, against the significant fitness scores. If strain JY304 has perfectly adapted to the post-stress state of the parental strain, we should see a shift of each of the points towards the x-axis, which is what we observe for genes with negative fitness scores (Figure 4.4, left half). This suggests that globally, genes that are essential for surviving acute lethal ethanol stress have adapted their expression levels to a post-stress state in strain JY304. In contrast, there was no significant change in expression fold change between positive fitness scores between Figure 3.7 and Figure 4.4. While certain functional categories in strain JY304, such as translation and ribosomal protein, move towards the post-stress state of the parental strain, in general, there is not a significant global shift in gene expression for those genes whose deletion helps survival.

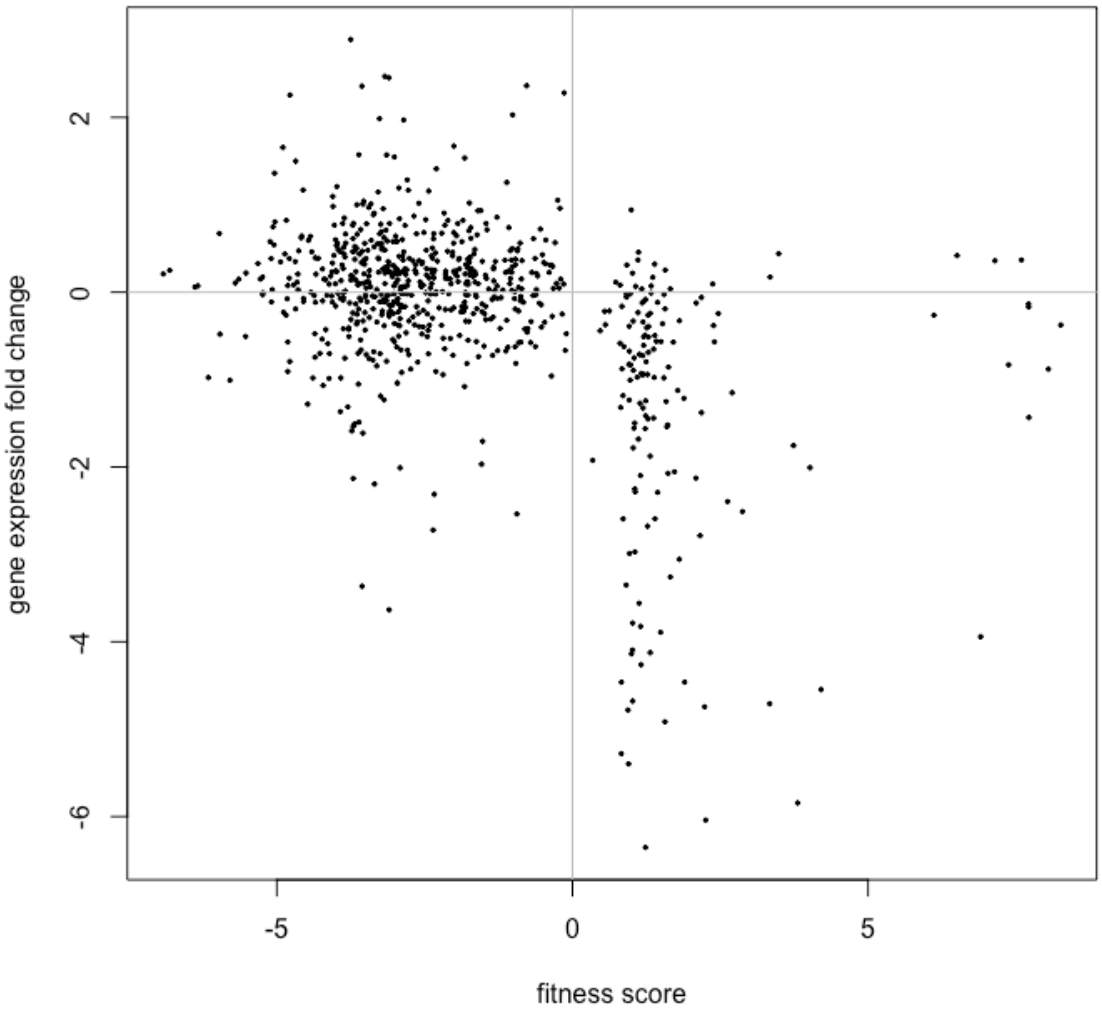


Figure 4.5 Gene expression fold change versus all significant fitness scores, using strain JY304 as the baseline pre-stress state
All significant fitness scores ($p<0.05$, Bonferroni corrected) were plotted against the ratio of their maximal post-stress expression fold change of the parental strain to the pre-stress expression level of strain JY304. The gene deletions with negative fitness scores (left side of graph) showed a shift in gene expression towards the x-axis ($p<1e-15$). The gene deletions with positive fitness scores (right side of graph) did not show any significant change in expression levels compared to the left half of Figure 3.7.

DISCUSSION

Here, we used laboratory evolution to evolve a wild-type strain with an initial 1% survival under lethal ethanol stress to more than 30% survival after 10 to 12 rounds of evolution. This

strain, which we named JY304, showed significant cross resistance to heat and hydrogen peroxide stress, indicating a general stress aspect, without compromising bulk growth rate. We determined that strain JY304 shifted its global expression level to that of the post-stress wild-type expression level, indicating a more stress-resistant state even during growth in optimal conditions.

In Zakrzewska et al., 2011, they were able to use the pooled yeast knockout library to determine the correlation between survival and growth rate of the gene deletions. Regardless of the type of stress they used, a 1% survival roughly corresponds to a growth rate of 0.34 doublings per hour, whereas a 30% survival roughly corresponds to a growth rate of 0.26 doublings per hour. In Lu C. et al., 2009, they forced their yeast to grow at several different rates in a chemostat. Testing survival at 50℃ for 5 minutes, they found that 1% survival roughly correspond to a growth rate of 0.44 doublings per hour, whereas a 30% survival roughly corresponds to a growth rate of 0.22 per hour. These studies demonstrated just how closely linked growth and stress defense are. Our study, however, showed that our evolved strain that has at least 30% survival had a bulk growth rate that was comparable to that of the wild-type strain.

In Figure 4.4, the decreased expression of genes that encode ribosomal proteins, ribosome biogenesis, and translation may suggest that bulk growth rate should be decreased. The pooled yeast deletion library used by Zakrzewska et al., 2011 did show that some genes with deletions in those three categories were slower growing. However, the growth rate of these strains was calculated by determining relative abundances of each strain within the entire pooled deletion library. Using only relative abundances is not as accurate as calculating growth rate of a clonal population. Additionally, the growth rate of each individual strain in the population will not necessarily be the same as if the individual strain were grown by itself, due to competition from all the other deletion strains. Welch et al., 2013 used 41 strains that had deletions in ribosome

biogenesis genes and showed that there was no correlation between growth rate and survival against desiccation stress. In fact, many of their ribosome biogenesis mutants showed greater than a 1000-fold increase in desiccation tolerance compared to wild-type but were actually in the quartile with the fastest growth rate.

While it may seem that we have evolved a strain that is able to optimize both growth rate and stress survival, the bulk growth rate of the population does not tell us whether individual cells actually have both optimized. It is certainly possible that strain JY304 shows a persistence or bet-hedging effect in which a small fraction of the population grows at a slow rate and is the main contributor of increased stress survival. Since we are only measuring bulk growth rate, we will not be able to truly determine if our population increases survival due to this subpopulation. We would need additional experiments to look at the distribution of growth rates within this population and determine if it is still the slow growing subpopulation that is the main contributor to increased survival. One aspect we were able to optimize, however, was recovery from stress. Whereas the initial rounds of evolution took three days to fully recover and saturate, we were able to evolve our strain to recover and reach saturation at two days or less.

Since ribosomal proteins, ribosome biogenesis, and translation all have decreased expression in strain JY304 even in optimal growth conditions, it seems that we have created a strain in which the repressed part of the ESR is mildly yet permanently activated. It has been shown that a decrease in repressed ESR transcripts causes a large increase in empty ribosomes, indicating that ribosomes are not limiting cell growth or translation, but instead directing translational capacity to induced mRNAs (Ho et al., 2018). A similar phenomenon has been proposed in bacterial cells as well, in which ribosomes are removed from the active pool upon nutrient starvation and osmotic stress to maintain elongation rates (Dai et al., 2017; Dai et al., 2018).

CONCLUSIONS

In this dissertation, we examined how yeast adapts to and survives acute exposure to lethal ethanol stress. Previous stress studies in ethanol mostly investigated the effects of mild doses of ethanol on yeast, with significant gene difference even between 8% and 11% ethanol (Kubota et al., 2004; Yoshikawa et al., 2009). While we found important pathways in our lethal stress studies that were consistent with many previous findings, we also discovered new insights into how yeast survives ethanol in the lethal range under acute exposure, experimental conditions that had not been well-studied in yeast.

We first investigated how wild-type haploid yeast responded to an acute exposure of threshold lethal ethanol stress. We showed that ribosome biogenesis, ribosomal proteins, and translation, all significantly downregulated categories of the environmental stress response, were also downregulated in response to threshold lethal ethanol stress. We also showed that heat shock proteins, which are significantly upregulated in the environmental stress response, are also upregulated in our experimental setup. Two functional categories, condensed chromosome and spore wall assembly, had never previously been described as having a role in ethanol stress survival. While genes in these categories contained a URS1 element that was associated with *HSF1*, this was still mostly in the context of meiosis and only demonstrated in diploid cells. The strong downregulation of genes containing the URS1 element as well as *IME1* suggests a possible adaptive role in response to threshold lethal ethanol stress, especially due to the fact that they have not previously been known to play in a role in haploid yeasts.

We next determined the fitness effect of every non-essential gene deletion under acute exposure to lethal ethanol stress. By calculating the fitness score of all gene deletions, we were able to determine which gene deletions help or hurt survival. At the positive and negative extremes of fitness scores, we saw genes that fall in similar functional categories, suggesting an enhanced

role for these pathways. Ho et al., 2018 proposed that this enhanced regulation may be used by the cell to shift the ethanol stress response away from being controlled transcriptionally normally to being controlled translationally instead. The anticorrelation between fitness scores and expression fold changes between the fitness profiling and transcriptional profiling for translation genes indicates a possible adaptive role for translational repression under acute exposure to lethal ethanol stress.

We next tried to improve stress survival to lethal ethanol stress by designing an experimental laboratory evolution protocol in which both stress defense as well as rapid growth of the bulk population was selected for. We were able to evolve a population of cells that showed a significant improvement in stress survival under lethal ethanol stress. Furthermore, the evolved strain showed cross protection to heat and hydrogen peroxide stress without compromising bulk growth rate. This evolved strain was found to be in a more stress-resistant state even when growing at optimal growth conditions.

A novel aspect of our experimental design is the very short amount of time of two minutes that yeast cells are exposed to ethanol. Most experiments performed on yeast are much longer, especially for studying acquired stress tolerance, in which the stress is there constantly until the cell mounts a transcriptional response to the stress. Here, the duration of two minutes does not give the cell time to mount a transcriptional response, which makes much of the cell's survival dependent on the pre-existing state prior to the onset of the stress. It is interesting to note how such a short duration affects the survival at a range of ethanol concentrations as seen in Figure 4.1C. For the wild-type strain, the range of ethanol concentrations from 20% to 25% decrease the survival in a dose-dependent manner over two orders of magnitude, but increasing the ethanol concentration to 26% suddenly decreases the survival another three orders of magnitude. A similar

behavior is seen with strain JY304. This evolved strain only demonstrates less than one magnitude decrease in survival between 20% and 25% ethanol but also shows a drastic drop in survival, of more than two orders of magnitude, between 25% to 26% ethanol. It would be interesting to further explore exactly what aspect of the stress defense mechanism is overcome by the ethanol stress at this extreme. Another interesting point to note is that both the wild-type strain and strain JY304 have slightly decreased survival at 20.5%. It seems that strain JY304 does not actually increase the threshold of lethality in regards to ethanol stress but rather just survives better than the wild-type strain within the lethal regime. An insightful future experiment would be to perform competition experiments between the wild-type strain and strain JY304 under various experimental conditions to determine how much more fit strain JY304 is compared to the wild-type strain.

While new insights can be gained from using such a short duration of stress, it remains to be seen whether these cells can survive longer term stresses. While we clearly demonstrated that strain JY304 survives better than the wild-type strain when stressed for two minutes under ethanol, heat, and hydrogen peroxide stress, we do not know whether it will fare as well when the duration of these stresses is longer. An interesting follow-up experiment would be to expose the wild-type and evolved strains to longer stress durations and track not only the survival but also gene expression changes.

A potential limitation of this study is the use of ethanol and pure water as the media in which the cells are stressed. Removing all nutrients, especially glucose, from the media ensures that the cells are not growing. Many stresses, especially starvation conditions, have demonstrated that the presence of glucose actually kills the cells faster than if no carbon source was present at all. This is due to the fact that cells that sense glucose may continue to try to divide, and this

utilization of valuable cellular resources that could have been used on stress defense actually kills the cells faster. To remove any possibility of this effect, we decided to remove all nutrients in the media before stressing the cells. The negative side effect of this is that the expression changes that result from this may be a combination of both ethanol stress and nutrient starvation. We do expect the effects to be predominantly from ethanol stress, however, because it is a stress at the threshold of lethality that the cell must quickly deal with in order to survive. In our experiments in which the cells were exposed to pure water only, they do not actually show any death for at least 24 hours (data not shown).

These results have potential applications in the food and biofuel industries. For example, many yeast strains used for winemaking will die at a certain ethanol concentration, limiting the alcohol content of the resulting wine. Insights from our evolved strain, which is more stress-resistant, could potentially be used to engineer yeast strains that can increase the alcohol content of wines before dying. Taken together, our results provide a key stepping stone in determining general principles that allow yeast and other microorganisms to adapt to and survive a diverse array of lethal stresses.

References

Alexandre, H., Ansanay-Galeote, V., Dequin, S., & Blondin, B. (2001, 05). Global gene expression during short-term ethanol stress in Saccharomyces cerevisiae. *FEBS Letters, 498*(1), 98-103. doi:10.1016/s0014-5793(01)02503-0

Ansanay-Galeote, V., Blondin, B., Dequin, S., & Sablayrolles, J. M. (2001). Stress effect of ethanol on fermentation kinetics by stationary-phase cells of Saccharomyces cerevisiae. *Biotechnology Letters*, *23*(9), 677-681.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Harris, M. A. (2000). Gene Ontology: tool for the unification of biology. *Nature genetics*, *25*(1), 25.

Ball, C. A., Dolinski, K., Dwight, S. S., Harris, M. A., Issel-Tarver, L., Kasarskis, A., ... & Kaloper, M. (2000). Integrating functional genomic information into the Saccharomyces genome database. *Nucleic acids research*, *28*(1), 77-80.

Bammert, G. F., & Fostel, J. M. (2000, 05). Genome-Wide Expression Patterns in Saccharomyces cerevisiae: Comparison of Drug Treatments and Genetic Alterations Affecting Biosynthesis of Ergosterol. *Antimicrobial Agents and Chemotherapy, 44*(5), 1255-1265. doi:10.1128/aac.44.5.1255-1265.2000

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., . . . Kim, J. F. (2009, 10). Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature, 461*(7268), 1243-1247. doi:10.1038/nature08480

Beaumont, H. J., Gallie, J., Kost, C., Ferguson, G. C., & Rainey, P. B. (2009, 11). Experimental evolution of bet hedging. *Nature, 462*(7269), 90-93. doi:10.1038/nature08504

Beck, T., & Hall, M. N. (1999, 12). The TOR signaling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature, 402*(6762), 689-692. doi:10.1038/45287

Bennett, A. F., Dao, K. M., & Lenski, R. E. (1990, 07). Rapid evolution in response to high-temperature selection. *Nature, 346*(6279), 79-81. doi:10.1038/346079a0

Betz, C., Schlenstedt, G., & Bailer, S. M. (2004, 04). Asr1p, a Novel Yeast Ring/PHD Finger Protein, Signals Alcohol Stress to the Nucleus. *Journal of Biological Chemistry, 279*(27), 28174-28181. doi:10.1074/jbc.m401595200

Blount, Z. D., Borland, C. Z., & Lenski, R. E. (2008, 06). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences, 105*(23), 7899-7906. doi:10.1073/pnas.0803151105

Boer, V. M., Amini, S., & Botstein, D. (2008, 05). Influence of genotype and nutrition on survival and metabolism of starving yeast. *Proceedings of the National Academy of Sciences, 105*(19), 6930-6935. doi:10.1073/pnas.0802601105

Bolger, A. M., Lohse, M., & Usadel, B. (2014, 04). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998, 01). Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast, 14*(2), 115-132. doi:10.1002/(sici)1097-0061(19980130)14:23.0.co;2-2

Brauer, M. J., Huttenhower, C., Airoldi, E. M., Rosenstein, R., Matese, J. C., Gresham, D., ... & Botstein, D. (2008). Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Molecular biology of the cell*, *19*(1), 352-367.

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016, 04). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology, 34*(5), 525-527. doi:10.1038/nbt.3519

Castrillo, J. I., Zeef, L. A., Hoyle, D. C., Zhang, N., Hayes, A., Gardner, D. C., ... & Rash, B. (2007). Growth control of the eukaryote cell: a systems biology study in yeast. *Journal of biology*, *6*(2), 4.

Causton, H. C., Ren, B., Koh, S. S., Harbison, C. T., Kanin, E., Jennings, E. G., . . . Young, R. A. (2001, 02). Remodeling of Yeast Genome Expression in Response to Environmental Changes. *Molecular Biology of the Cell, 12*(2), 323-337. doi:10.1091/mbc.12.2.323

Chandler, M., Stanley, G. A., Rogers, P., & Chambers, P. (2004). A genomic approach to defining the ethanol stress response in the yeast Saccharomyces cerevisiae. *Ann Microbiol*, *54*(4), 427-454.

Charoenbhakdi, S., Dokpikul, T., Burphan, T., Techo, T., & Auesukaree, C. (2016, 03). Vacuolar H -ATPase Protects Saccharomyces cerevisiae Cells against Ethanol-Induced Oxidative and Cell Wall Stresses. *Applied and Environmental Microbiology, 82*(10), 3121-3130. doi:10.1128/aem.00376-16

Choder, M., & Young, R. A. (1993, 11). A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. *Molecular and Cellular Biology, 13*(11), 6984-6991. doi:10.1128/mcb.13.11.6984

Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO, Herskowitz I. The transcriptional program of sporulation in budding yeast. Science. 1998 Oct 23;282(5389):699-705.

Cooper, V. S., & Lenski, R. E. (2000, 10). The population genetics of ecological specialization in evolving Escherichia coli populations. *Nature, 407*(6805), 736-739. doi:10.1038/35037572

Cowart, L. A., & Obeid, L. M. (2007, 03). Yeast sphingolipids: Recent developments in understanding biosynthesis, regulation, and function. *Biochimica Et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 1771*(3), 421-431. doi:10.1016/j.bbalip.2006.08.005

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., ... & Hwa, T. (2017). Reduction of translating ribosomes enables Escherichia coli to maintain elongation rates during slow growth. Nature microbiology, 2(2), 16231.

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., ... & Hwa, T. (2018). Slowdown of translational elongation in Escherichia coli under hyperosmotic stress. MBio, 9(1), e02375-17.

D'amore, T., Panchal, C. J., Russell, I., & Stewart, G. G. (1989, 01). A Study of Ethanol Tolerance in Yeast. *Critical Reviews in Biotechnology, 9*(4), 287-304. doi:10.3109/07388558909036740

Deatherage, D. E., & Barrick, J. E. (2014). Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using breseq. *Methods in Molecular Biology Engineering and Analyzing Multicellular Systems,* 165-188. doi:10.1007/978-1-4939-0554-6_12

Dragosits, M., Mozhayskiy, V., Quinones-Soto, S., Park, J., & Tagkopoulos, I. (2014, 04). Evolutionary potential, cross-stress behavior and the genetic basis of acquired stress resistance in Escherichia coli. *Molecular Systems Biology, 9*(1), 643-643. doi:10.1038/msb.2012.76

Dragosits, M., & Mattanovich, D. (2013). Adaptive laboratory evolution – principles and applications for biotechnology. *Microbial Cell Factories, 12*(1), 64. doi:10.1186/1475-2859-12-64

Erkine, A. M., Magrogan, S. F., Sekinger, E. A., & Gross, D. S. (1999). Cooperative binding of heat shock factor to the yeast HSP82 promoter in vivo and in vitro. Molecular and cellular biology, 19(3), 1627-1639.

Fujita, K., Matsuyama, A., Kobayashi, Y., & Iwahashi, H. (2004, 07). Comprehensive gene expression analysis of the response to straight-chain alcohols in Saccharomyces cerevisiae using cDNA microarray. *Journal of Applied Microbiology, 97*(1), 57-67. doi:10.1111/j.1365-2672.2004.02290.x

Fujita, K., Matsuyama, A., Kobayashi, Y., & Iwahashi, H. (2006, 08). The genome-wide screening of yeast deletion mutants to identify the genes required for tolerance to ethanol

and other alcohols. *FEMS Yeast Research, 6*(5), 744-750. doi:10.1111/j.1567-1364.2006.00040.x

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., . . . Brown, P. O. (2000, 12). Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. *Molecular Biology of the Cell, 11*(12), 4241-4257. doi:10.1091/mbc.11.12.4241

Gasch, A. P., Huang, M., Metzner, S., Botstein, D., Elledge, S. J., & Brown, P. O. (2001, 10). Genomic Expression Responses to DNA-damaging Agents and the Regulatory Role of the Yeast ATR Homolog Mec1p. *Molecular Biology of the Cell, 12*(10), 2987-3003. doi:10.1091/mbc.12.10.2987

Gasch, A. P., & Werner-Washburne, M. (2002, 09). The genomics of yeast responses to environmental stress and starvation. *Functional & Integrative Genomics, 2*(4-5), 181-192. doi:10.1007/s10142-002-0058-2

Gasch, A. P. (2003). The environmental stress response: a common yeast response to diverse environmental stresses. In *Yeast stress responses* (pp. 11-70). Springer, Berlin, Heidelberg.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., ... & Arkin, A. P. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *nature*, *418*(6896), 387.

Gietz, R. D., & Woods, R. A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Guide to Yeast Genetics and Molecular and Cell Biology - Part B Methods in Enzymology,* 87-96. doi:10.1016/s0076-6879(02)50957-5

Goodarzi, H., Elemento, O., & Tavazoie, S. (2009, 12). Revealing Global Regulatory Perturbations across Human Cancers. *Molecular Cell, 36*(5), 900-911. doi:10.1016/j.molcel.2009.11.016

Gross, C., Kelleher, M., Iyer, V. R., Brown, P. O., & Winge, D. R. (2000). Identification of the copper regulon in Saccharomyces cerevisiae by DNA microarrays. *Journal of biological chemistry*, *275*(41), 32310-32316.

Hallsworth, J. E., Nomura, Y., & Iwahara, M. (1998, 01). Ethanol-induced water stress and fungal growth. *Journal of Fermentation and Bioengineering, 86*(5), 451-456. doi:10.1016/s0922-338x(98)80150-5

Han, T. X., Xu, X., Zhang, M., Peng, X., & Du, L. (2010). Global fitness profiling of fission yeast deletion strains by barcode sequencing. *Genome Biology, 11*(6). doi:10.1186/gb-2010-11-6-r60

Hillenmeyer, M. E., Fung, E., Wildenhain, J., Pierce, S. E., Hoon, S., Lee, W., . . . Giaever, G. (2008, 04). The Chemical Genomic Portrait of Yeast: Uncovering a Phenotype for All Genes. *Science, 320*(5874), 362-365. doi:10.1126/science.1150021

Ho, Y., Shishkova, E., Hose, J., Coon, J. J., & Gasch, A. P. (2018, 08). Decoupling Yeast Cell Division and Stress Defense Implicates mRNA Repression in Translational Reallocation during Stress. *Current Biology, 28*(16). doi:10.1016/j.cub.2018.06.044

Ho, Y., & Gasch, A. P. (2015, 05). Exploiting the yeast stress-activated signaling network to inform on stress biology and disease signaling. *Current Genetics, 61*(4), 503-511. doi:10.1007/s00294-015-0491-0

Hottiger, T., Boller, T., & Wiemken, A. (1987, 08). Rapid changes of heat and desiccation tolerance correlated with changes of trehalose content in Saccharomyces cerevisiae cells subjected to temperature shifts. *FEBS Letters, 220*(1), 113-115. doi:10.1016/0014-5793(87)80886-4

Hu, X. H., Wang, M. H., Tan, T., Li, J. R., Yang, H., Leach, L., . . . Luo, Z. W. (2006, 12). Genetic Dissection of Ethanol Tolerance in the Budding Yeast Saccharomyces cerevisiae. *Genetics, 175*(3), 1479-1487. doi:10.1534/genetics.106.065292

Hughes, B. S., Cullum, A. J., & Bennett, A. F. (2007, 07). Evolutionary Adaptation to Environmental pH in Experimental Lineages of Escherichia Coli. *Evolution, 61*(7), 1725-1734. doi:10.1111/j.1558-5646.2007.00139.x

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., . . . Friend, S. H. (2000, 07). Functional Discovery via a Compendium of Expression Profiles. *Cell, 102*(1), 109-126. doi:10.1016/s0092-8674(00)00015-5

Izawa, S., Ikeda, K., Kita, T., & Inoue, Y. (2006, 01). Asr1, an alcohol-responsive factor of Saccharomyces cerevisiae, is dispensable for alcoholic fermentation. *Applied Microbiology and Biotechnology, 72*(3), 560-565. doi:10.1007/s00253-005-0294-1

Jaskowiak, P. A., Campello, R. J., & Costa, I. G. (2014). On the selection of appropriate distances for gene expression data clustering. *BMC Bioinformatics, 15*(Suppl 2). doi:10.1186/1471-2105-15-s2-s2

Kao, K. C., & Sherlock, G. (2008, 11). Molecular characterization of clonal interference during adaptive evolution in asexual populations of Saccharomyces cerevisiae. *Nature Genetics, 40*(12), 1499-1504. doi:10.1038/ng.280

Kobayashi, N., & Mcentee, K. (1990, 09). Evidence for a heat shock transcription factor-independent mechanism for heat shock induction of transcription in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences, 87*(17), 6550-6554. doi:10.1073/pnas.87.17.6550

Kobayashi, N., & Mcentee, K. (1993, 01). Identification of cis and trans components of a novel heat shock stress regulatory pathway in Saccharomyces cerevisiae. *Molecular and Cellular Biology, 13*(1), 248-256. doi:10.1128/mcb.13.1.248

Kubota, S., Takeo, I., Kume, K., Kanai, M., Shitamukai, A., Mizunuma, M., . . . Hirata, D. (2004, 01). Effect of Ethanol on Cell Growth of Budding Yeast: Genes That Are Important for Cell Growth in the Presence of Ethanol. *Bioscience, Biotechnology, and Biochemistry, 68*(4), 968-972. doi:10.1271/bbb.68.968

Kuge, S., Jones, N., & Nomoto, A. (1997). Regulation of yAP-1 nuclear localization in response to oxidative stress. *The EMBO journal*, *16*(7), 1710-1720.

Kupiec M, Byers B, Esposito RE, Mitchell AP. The molecular and cellular biology of the yeast Saccharomyces. Cell Cycle and Cell Biology. 1997;3.

Kwast, K. E., Lai, L., Menda, N., James, D. T., Aref, S., & Burke, P. V. (2002, 01). Genomic Analyses of Anaerobically Induced Genes in Saccharomyces cerevisiae: Functional Roles of Rox1 and Other Factors in Mediating the Anoxic Response. *Journal of Bacteriology, 184*(1), 250-265. doi:10.1128/jb.184.1.250-265.2002

Langmead, B., & Salzberg, S. L. (2012, 03). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

Lee, S., Pellicioli, A., Demeter, J., Vaze, M., Gasch, A., Malkova, A., . . . Haber, J. (2000, 01). Arrest, Adaptation, and Recovery following a Chromosome Double-strand Break in Saccharomyces cerevisiae. *Cold Spring Harbor Symposia on Quantitative Biology, 65*(0), 303-314. doi:10.1101/sqb.2000.65.303

Lenski, R. E., Mongold, J. A., Sniegowski, P. D., Travisano, M., Vasi, F., Gerrish, P. J., & Schmidt, T. M. (1998). Evolution of competitive fitness in experimental populations of E. coli: what makes one genotype a better competitor than another?. *Antonie van Leeuwenhoek*, *73*(1), 35-47.

Levy, S. F., Ziv, N., & Siegal, M. L. (2012, 05). Bet Hedging in Yeast by Heterogeneous, Age-Correlated Expression of a Stress Protectant. *PLoS Biology, 10*(5). doi:10.1371/journal.pbio.1001325

Leão, C., & Uden, N. V. (1984, 07). Effects of ethanol and other alkanols on passive proton influx in the yeast Saccharomyces cerevisiae. *Biochimica Et Biophysica Acta (BBA) - Biomembranes, 774*(1), 43-48. doi:10.1016/0005-2736(84)90272-4

Lindquist, S., & Craig, E. A. (1988, 12). The Heat-Shock Proteins. *Annual Review of Genetics, 22*(1), 631-677. doi:10.1146/annurev.ge.22.120188.003215

Longo, V., Shadel, G., Kaeberlein, M., & Kennedy, B. (2012, 07). Replicative and Chronological Aging in Saccharomyces cerevisiae. *Cell Metabolism, 16*(1), 18-31. doi:10.1016/j.cmet.2012.06.002

Love, M. I., Huber, W., & Anders, S. (2014, 12). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*(12). doi:10.1186/s13059-014-0550-8

Lu, C., Brauer, M. J., & Botstein, D. (2009, 02). Slow Growth Induces Heat-Shock Resistance in Normal and Respiratory-deficient Yeast. *Molecular Biology of the Cell, 20*(3), 891-903. doi:10.1091/mbc.e08-08-0852

Lyons, T. J., Gasch, A. P., Gaither, L. A., Botstein, D., Brown, P. O., & Eide, D. J. (2000, 07). Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. *Proceedings of the National Academy of Sciences, 97*(14), 7957-7962. doi:10.1073/pnas.97.14.7957

Mager, W. H., & De Kruijff, A. J. (1995). Stress-induced transcriptional activation. *Microbiological reviews*, *59*(3), 506-531.

Mager, W. H., & Ferreira, P. M. (1993, 02). Stress response of yeast. *Biochemical Journal, 290*(1), 1-13. doi:10.1042/bj2900001

Marchler, G., Schüller, C., Adam, G., & Ruis, H. (1993, 05). A Saccharomyces cerevisiae UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions. *The EMBO Journal, 12*(5), 1997-2003. doi:10.1002/j.1460-2075.1993.tb05849.x

Martin, M. (2011, 05). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal, 17*(1), 10. doi:10.14806/ej.17.1.200

Martínez-Pastor, M. T., Marchler, G., Schüller, C., Marchler-Bauer, A., Ruis, H., & Estruch, F. (1996, 05). The Saccharomyces cerevisiae zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *The EMBO Journal, 15*(9), 2227-2235. doi:10.1002/j.1460-2075.1996.tb00576.x

Meaden, P. G., Arneborg, N., Guldfeldt, L. U., Siegumfeldt, H., & Jakobsen, M. (1999, 09). Endocytosis and vacuolar morphology in Saccharomyces cerevisiae are altered in response to ethanol stress or heat shock. *Yeast, 15*(12), 1211-1222. doi:10.1002/(sici)1097-0061(19990915)15:123.0.co;2-h

Mewes, H. W., Frishman, D., Gruber, C., Geier, B., Haase, D., Kaps, A., ... & Stocker, S. (2000). MIPS: a database for genomes and protein sequences. *Nucleic acids research*, *28*(1), 37-40.

Mishra, P., & Prasad, R. (1989, 03). Relationship between ethanol tolerance and fatty acyl composition of Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology, 30*(3). doi:10.1007/bf00256221

Momose, Y., & Iwahashi, H. (2001). Bioassay of Cadmium Using A DNA Microarray: Genome-Wide Expression Patterns of Saccharomyces Cerevisiae Response to Cadmium. *Environmental Toxicology and Chemistry, 20*(10), 2353. doi:10.1897/1551-5028(2001)0202.0.co;2

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., & Snyder, M. (2008, 06). The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing. *Science, 320*(5881), 1344-1349. doi:10.1126/science.1158441

Nasmyth KA. Molecular genetics of yeast mating type. Annual review of genetics. 1982 Dec;16(1):439-500.

Navarro-Tapia, E., Nana, R. K., Querol, A., & Pérez-Torrado, R. (2016, 02). Ethanol Cellular Defense Induce Unfolded Protein Response in Yeast. *Frontiers in Microbiology, 7*. doi:10.3389/fmicb.2016.00189

Ogawa, N., Derisi, J., & Brown, P. O. (2000, 12). New Components of a System for Phosphate Accumulation and Polyphosphate Metabolism in Saccharomyces cerevisiae Revealed by Genomic Expression Analysis. *Molecular Biology of the Cell, 11*(12), 4309-4321. doi:10.1091/mbc.11.12.4309

Pan, Y., Schroeder, E., Ocampo, A., Barrientos, A., & Shadel, G. (2011, 06). Regulation of Yeast Chronological Life Span by TORC1 via Adaptive Mitochondrial ROS Signaling. *Cell Metabolism, 13*(6), 668-678. doi:10.1016/j.cmet.2011.03.018

Paquin, C., & Adams, J. (1983, 04). Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature, 302*(5908), 495-500. doi:10.1038/302495a0

Paquin, C. E., & Adams, J. (1983, 11). Relative fitness can decrease in evolving asexual populations of S. cerevisiae. *Nature, 306*(5941), 368-371. doi:10.1038/306368a0

Piper, P. W., Talreja, K., Panaretou, B., Moradas-Ferreira, P., Byrne, K., Praekelt, U. M., . . . Boucherie, H. (1994, 11). Induction of major heat-shock proteins of Saccharomyces cerevisiae, including plasma membrane Hsp30, by ethanol levels above a critical threshold. *Microbiology, 140*(11), 3031-3038. doi:10.1099/13500872-140-11-3031

Piper, P. (1993, 08). Molecular events associated with acquisition of heat tolerance by the yeast Saccharomyces cerevisiae. *FEMS Microbiology Reviews, 11*(4), 339-355. doi:10.1016/0168-6445(93)90005-t

Powers, R. W., Kaeberlein, M., Caldwell, S. D., Kennedy, B. K., & Fields, S. (2006). Extension

of chronological life span in yeast by decreased TOR pathway signaling. *Genes & development*, *20*(2), 174-184.

Rep, M., Krantz, M., Thevelein, J. M., & Hohmann, S. (2000). The transcriptional response of Saccharomyces cerevisiae to osmotic shock Hot1p and Msn2p/Msn4p are required for the induction of subsets of high osmolarity glycerol pathway-dependent genes. *Journal of Biological Chemistry*, *275*(12), 8290-8300.

Rep, M., Reiser, V., Gartner, U., Thevelein, J. M., Hohmann, S., Ammerer, G., & Ruis, H. (1999, 08). Osmotic Stress-Induced Gene Expression in Saccharomyces cerevisiae Requires Msn1p and the Novel Nuclear Factor Hot1p. *Molecular and Cellular Biology, 19*(8), 5474-5485. doi:10.1128/mcb.19.8.5474

Roy, N. C., Altermann, E., Park, Z. A., & Mcnabb, W. C. (2011, 03). A comparison of analog and Next-Generation transcriptomic tools for mammalian studies. *Briefings in Functional Genomics, 10*(3), 135-150. doi:10.1093/bfgp/elr005

Rubin-Bejerano I, Mandel S, Robzyk K, Kassir Y. Induction of meiosis in Saccharomyces cerevisiae depends on conversion of the transcriptional represssor Ume6 to a positive regulator by its regulated association with the transcriptional activator Ime1. Molecular and cellular biology. 1996 May 1;16(5):2518-26.

Ruis, H., & Schüller, C. (1995, 11). Stress signaling in yeast. *BioEssays, 17*(11), 959-965. doi:10.1002/bies.950171109

Ryan, O. W., Poddar, S., & Cate, J. H. (2016, 06). CRISPR–Cas9 Genome Engineering in Saccharomyces cerevisiae Cells. *Cold Spring Harbor Protocols, 2016*(6). doi:10.1101/pdb.prot086827

Ryan, O. W., & Cate, J. H. (2014). Multiplex Engineering of Industrial Yeast Genomes Using CRISPRm. *Methods in Enzymology The Use of CRISPR/Cas9, ZFNs, and TALENs in Generating Site-Specific Genome Alterations,* 473-489. doi:10.1016/b978-0-12-801185-0.00023-4

Sanchez, Y., Taulien, J., Borkovich, K., & Lindquist, S. (1992, 06). Hsp104 is required for tolerance to many forms of stress. *The EMBO Journal, 11*(6), 2357-2364. doi:10.1002/j.1460-2075.1992.tb05295.x

Sanctis, V. D., Bertozzi, C., Costanzo, G., Mauro, E. D., & Negri, R. (2001, 10). Cell Cycle Arrest Determines the Intensity of the Global Transcriptional Response of Saccharomyces cerevisiae to Ionizing Radiation. *Radiation Research, 156*(4), 379-387. doi:10.1667/0033-7587(2001)156[0379:ccadti]2.0.co;2

Schena, M., Shalon, D., Davis, R. W., & Brown, P. O. (1995, 10). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science, 270*(5235), 467-470. doi:10.1126/science.270.5235.467

Schüller, C., Brewster, J., Alexander, M., Gustin, M., & Ruis, H. (1994, 09). The HOG pathway controls osmotic regulation of transcription via the stress response element (STRE) of the Saccharomyces cerevisiae CTT1 gene. *The EMBO Journal, 13*(18), 4382-4389. doi:10.1002/j.1460-2075.1994.tb06758.x

Smith, A., Ward, M. P., & Garrett, S. (1998). Yeast PKA represses Msn2p/Msn4p-dependent gene expression to regulate growth, stress response and glycogen accumulation. *The EMBO journal*, *17*(13), 3556-3564.

Sniegowski, P. D., Gerrish, P. J., & Lenski, R. E. (1997, 06). Evolution of high mutation rates in experimental populations of E. coli. *Nature, 387*(6634), 703-705. doi:10.1038/42701

Spies, D., & Ciaudo, C. (2015). Dynamics in Transcriptomics: Advancements in RNA-seq Time Course and Downstream Analysis. *Computational and Structural Biotechnology Journal, 13*, 469-477. doi:10.1016/j.csbj.2015.08.004

Stanley, D., Bandara, A., Fraser, S., Chambers, P. J., & Stanley, G. A. (2009, 12). The ethanol stress response and ethanol tolerance of Saccharomyces cerevisiae. *Journal of Applied Microbiology*. doi:10.1111/j.1364-5072.2009.04657.x

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011, 07). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS ONE, 6*(7). doi:10.1371/journal.pone.0021800

Swan, T. M., & Watson, K. (1999). Stress tolerance in a yeast lipid mutant: Membrane lipids influence tolerance to heat and ethanol independently of heat shock proteins and trehalose. *Canadian Journal of Microbiology, 45*(6), 472-479. doi:10.1139/cjm-45-6-472

Szent-Gyorgyi, C. (1995). A bipartite operator interacts with a heat shock element to mediate early meiotic induction of Saccharomyces cerevisiae HSP82. Molecular and cellular biology, 15(12), 6754-6769.

Tagkopoulos, I., Liu, Y., & Tavazoie, S. (2008, 06). Predictive Behavior Within Microbial Genetic Networks. *Science, 320*(5881), 1313-1317. doi:10.1126/science.1154456

Takemura, R., Inoue, Y., & Izawa, S. (2004). Stress response in yeast mRNA export factor: reversible changes in Rat8p localization are caused by ethanol stress but not heat shock. *Journal of cell science*, *117*(18), 4189-4197.

Treger, J. M., Schmitt, A. P., Simon, J. R., & Mcentee, K. (1998, 10). Transcriptional Factor Mutations Reveal Regulatory Complexities of Heat Shock and Newly Identified Stress Genes in Saccharomyces cerevisiae. *Journal of Biological Chemistry, 273*(41), 26875-26879. doi:10.1074/jbc.273.41.26875

Tremblay, P., Summers, Z. M., Glaven, R. H., Nevin, K. P., Zengler, K., Barrett, C. L., . . . Lovley, D. R. (2010, 07). A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in Geobacter sulfurreducens revealed by adaptive evolution. *Environmental Microbiology, 13*(1), 13-23. doi:10.1111/j.1462-2920.2010.02302.x

Uffenbeck, S. R., & Krebs, J. E. (2006). The role of chromatin structure in regulating stress-induced transcription in Saccharomyces cerevisiae. Biochemistry and cell biology, 84(4), 477-489.

Voorst, F. V., Houghton-Larsen, J., Jønson, L., Kielland-Brandt, M. C., & Brandt, A. (2006). Genome-wide identification of genes required for growth of Saccharomyces cerevisiae under ethanol stress. *Yeast, 23*(5), 351-359. doi:10.1002/yea.1359

Wagstaff, J. E., Klapholz, S., & Esposito, R. E. (1982, 05). Meiosis in haploid yeast. *Proceedings of the National Academy of Sciences, 79*(9), 2986-2990. doi:10.1073/pnas.79.9.2986

Walker, G. M. (2000). *Yeast physiology and biotechnology*. John Wiley & Sons.

Wang, Z., Gerstein, M., & Snyder, M. (2009, 01). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics, 10*(1), 57-63. doi:10.1038/nrg2484

Warner, J. R. (1999, 11). The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences, 24*(11), 437-440. doi:10.1016/s0968-0004(99)01460-7

Wei, M., Fabrizio, P., Madia, F., Hu, J., Ge, H., Li, L. M., & Longo, V. D. (2009, 05). Tor1/Sch9-Regulated Carbon Source Substitution Is as Effective as Calorie Restriction in Life Span Extension. *PLoS Genetics, 5*(5). doi:10.1371/journal.pgen.1000467

Welch, A. Z., Gibney, P. A., Botstein, D., & Koshland, D. E. (2013, 01). TOR and RAS pathways regulate desiccation tolerance in Saccharomyces cerevisiae. *Molecular Biology of the Cell, 24*(2), 115-128. doi:10.1091/mbc.e12-07-0524

Wemmie, J. A., Steggerda, S. M., & Moye-Rowley, W. S. (1997, 03). The Saccharomyces cerevisiae AP-1 Protein Discriminates between Oxidative Stress Elicited by the Oxidants H2O2and Diamide. *Journal of Biological Chemistry, 272*(12), 7908-7914. doi:10.1074/jbc.272.12.7908

Wieser, R., Adam, G., Wagner, A., Schüller, C., Marchler, G., Ruis, H., ... & Bilinski, T. (1991). Heat shock factor-independent heat control of transcription of the CTT1 gene encoding the cytosolic catalase T of Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *266*(19), 12406-12411.

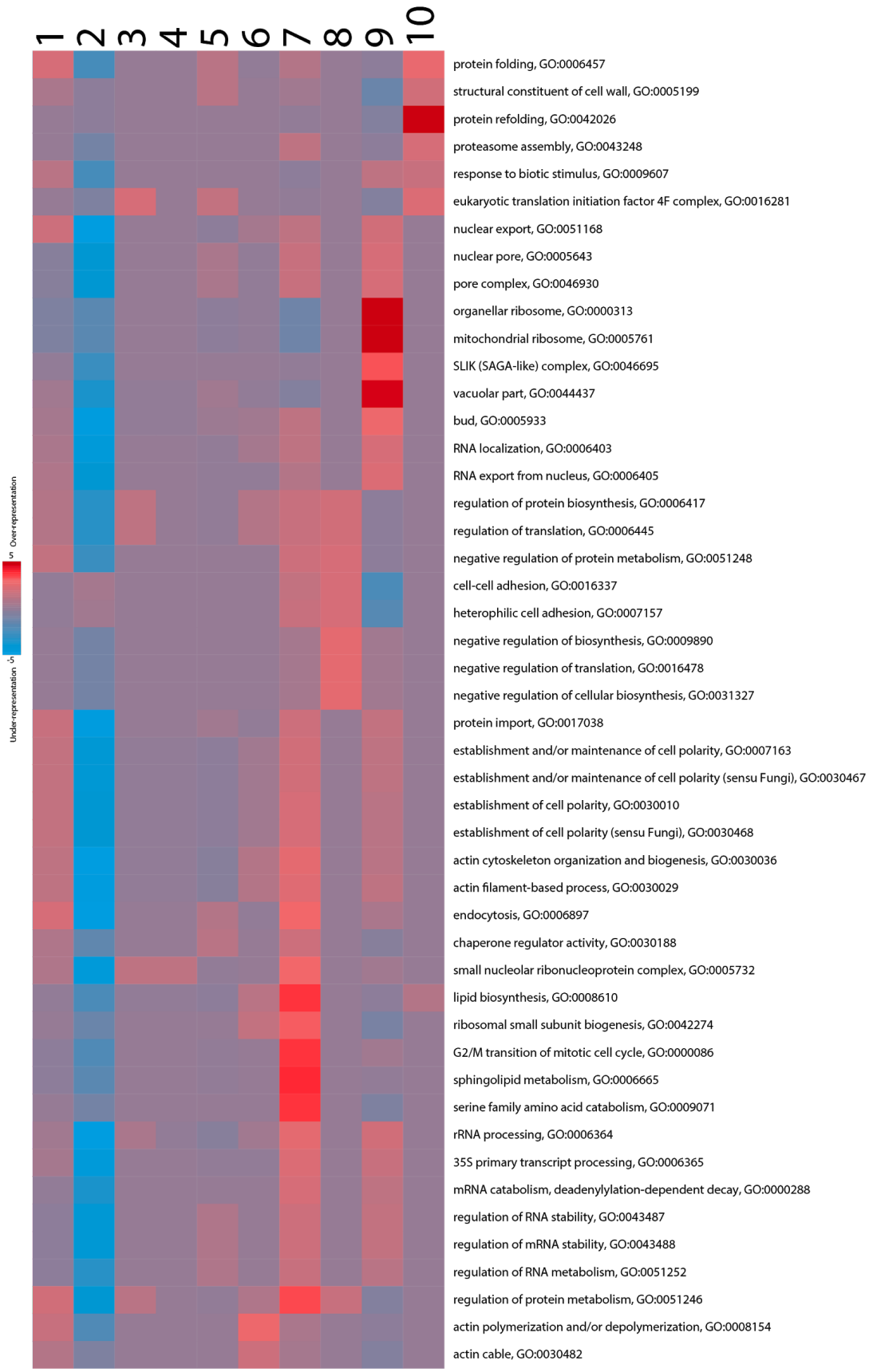
Yaakov, G., Lerner, D., Bentele, K., Steinberger, J., & Barkai, N. (2017, 01). Coupling phenotypic persistence to DNA damage increases genetic diversity in severe stress. *Nature Ecology & Evolution, 1*(1), 0016. doi:10.1038/s41559-016-0016

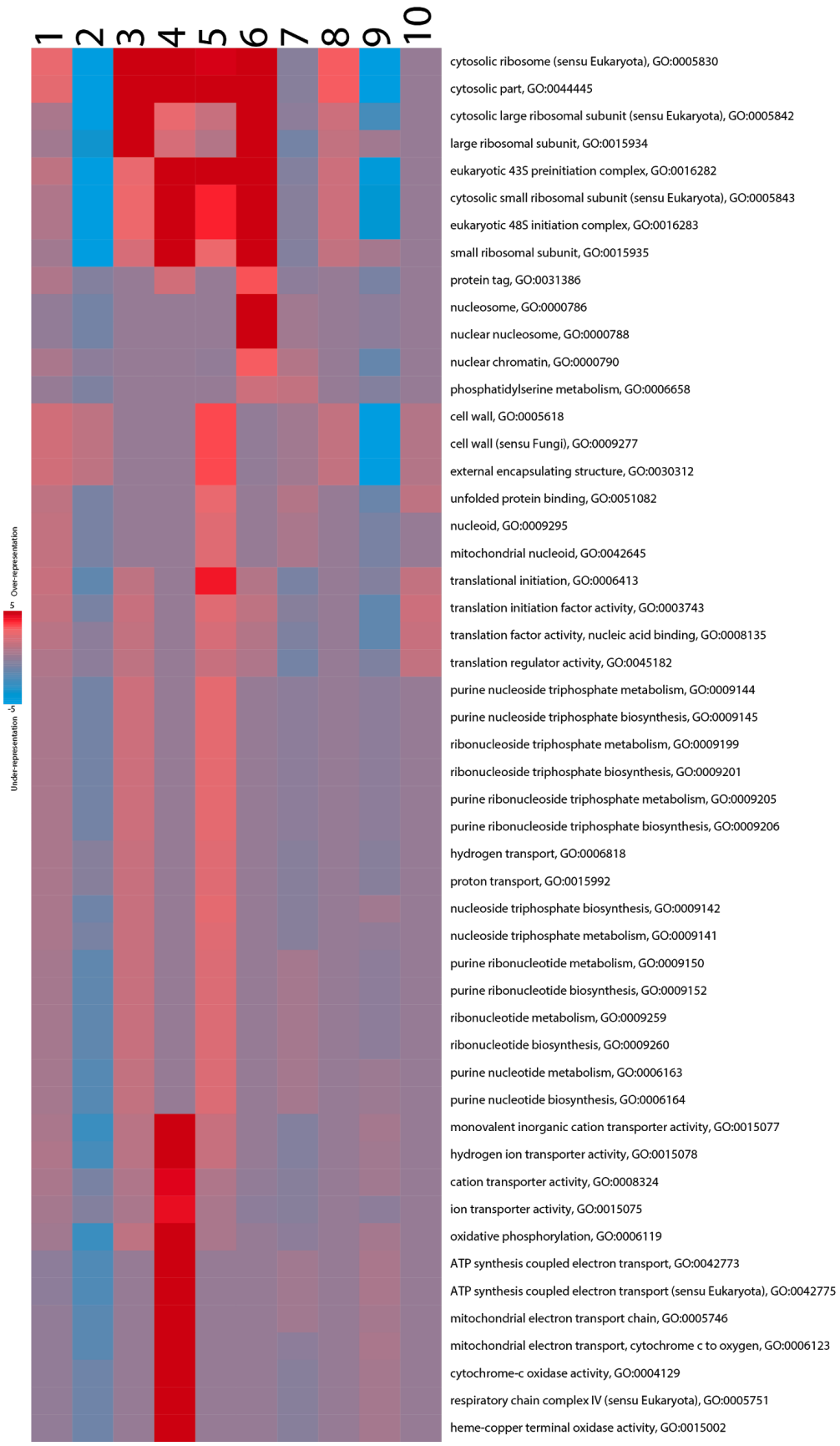
Yang, H., Ren, Q., & Zhang, Z. (2006, 12). Chromosome or chromatin condensation leads to meiosis or apoptosis in stationary yeast (Saccharomyces cerevisiae) cells. *FEMS Yeast Research, 6*(8), 1254-1263. doi:10.1111/j.1567-1364.2006.00123.x

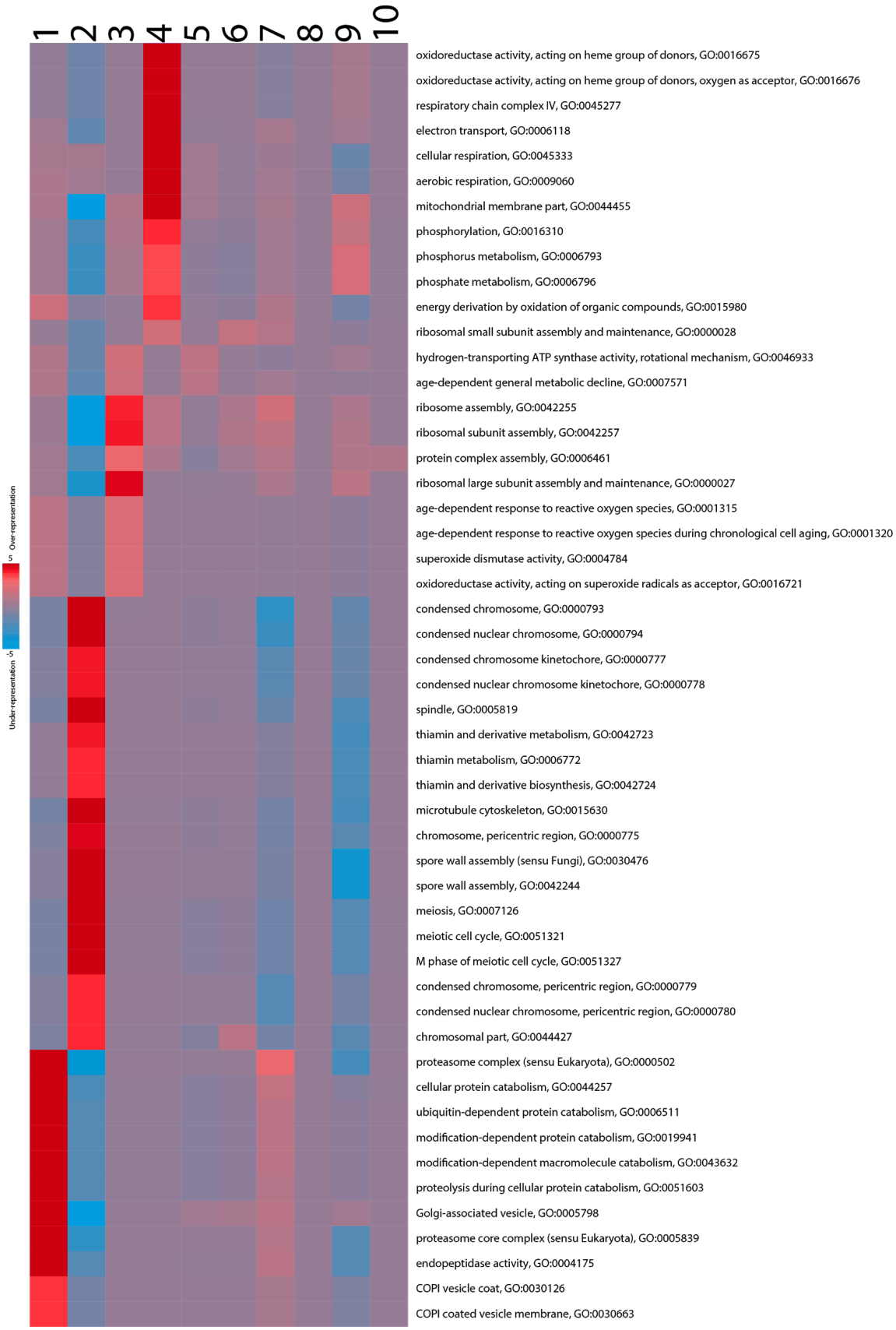
Yoshikawa, K., Tanaka, T., Furusawa, C., Nagahisa, K., Hirasawa, T., & Shimizu, H. (2009, 02). Comprehensive phenotypic analysis for identification of genes affecting growth under ethanol stress in Saccharomyces cerevisiae. *FEMS Yeast Research, 9*(1), 32-44. doi:10.1111/j.1567-1364.2008.00456.x

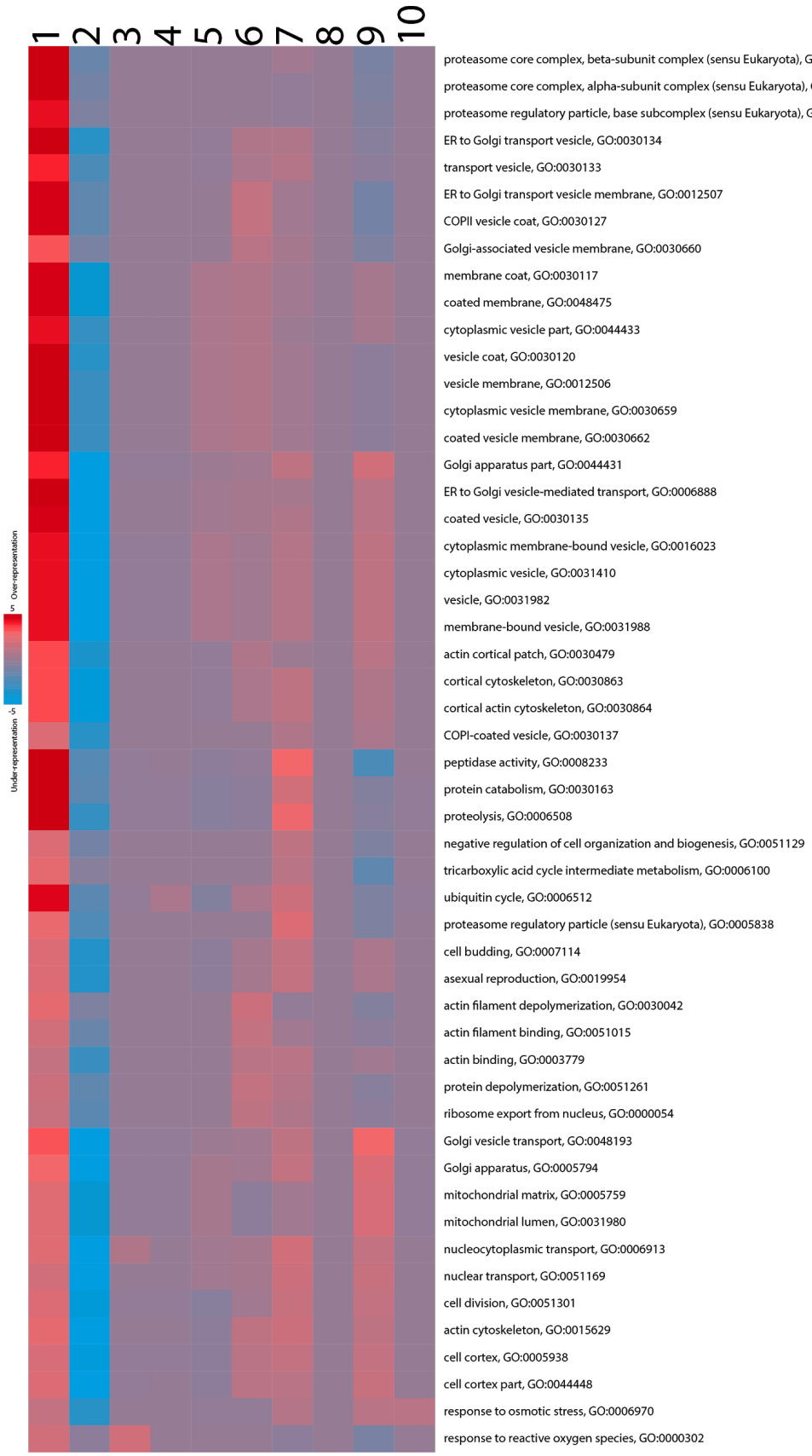
Zakrzewska, A., Eikenhorst, G. V., Burggraaff, J. E., Vis, D. J., Hoefsloot, H., Delneri, D., . . . Smits, G. J. (2011, 11). Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. *Molecular Biology of the Cell, 22*(22), 4435-4446. doi:10.1091/mbc.e10-08-0721

APPENDIX A: SUPPLEMENTARY FIGURES AND TABLES


1
2
3
4
5
6
7
8
9
10
protein folding, GO:0006457
structural constituent of cell wall, GO:0005199
protein refolding, GO:0042026
proteasome assembly, GO:0043248
response to biotic stimulus, GO:0009607
eukaryotic translation initiation factor 4F complex, GO:0016281
nuclear export, GO:0051168
nuclear pore, GO:0005643
pore complex, GO:0046930
organellar ribosome, GO:0000313
mitochondrial ribosome, GO:0005761
SLIK (SAGA-like) complex, GO:0046695
vacuolar part, GO:0044437
bud, GO:0005933
RNA localization, GO:0006403
RNA export from nucleus, GO:0006405
regulation of protein biosynthesis, GO:0006417
regulation of translation, GO:0006445
negative regulation of protein metabolism, GO:0051248
cell-cell adhesion, GO:0016337
heterophilic cell adhesion, GO:0007157
negative regulation of biosynthesis, GO:0009890
negative regulation of translation, GO:0016478
negative regulation of cellular biosynthesis, GO:0031327
protein import, GO:0017038
establishment and/or maintenance of cell polarity, GO:0007163
establishment and/or maintenance of cell polarity (sensu Fungi), GO:0030467
establishment of cell polarity, GO:0030010
establishment of cell polarity (sensu Fungi), GO:0030468
actin cytoskeleton organization and biogenesis, GO:0030036
actin filament-based process, GO:0030029
endocytosis, GO:0006897
chaperone regulator activity, GO:0030188
small nucleolar ribonucleoprotein complex, GO:0005732
lipid biosynthesis, GO:0008610
ribosomal small subunit biogenesis, GO:0042274
G2/M transition of mitotic cell cycle, GO:0000086
sphingolipid metabolism, GO:0006665
serine family amino acid catabolism, GO:0009071
rRNA processing, GO:0006364
35S primary transcript processing, GO:0006365
mRNA catabolism, deadenylylation-dependent decay, GO:0000288
regulation of RNA stability, GO:0043487
regulation of mRNA stability, GO:0043488
regulation of RNA metabolism, GO:0051252
regulation of protein metabolism, GO:0051246
actin polymerization and/or depolymerization, GO:0008154
actin cable, GO:0030482
Over-representation
5
-5
Under-representation


1
2
3
4
5
6
7
8
9
10
Over-representation
5
-5
Under-representation
cytosolic ribosome (sensu Eukaryota), GO:0005830
cytosolic part, GO:0044445
cytosolic large ribosomal subunit (sensu Eukaryota), GO:0005842
large ribosomal subunit, GO:0015934
eukaryotic 43S preinitiation complex, GO:0016282
cytosolic small ribosomal subunit (sensu Eukaryota), GO:0005843
eukaryotic 48S initiation complex, GO:0016283
small ribosomal subunit, GO:0015935
protein tag, GO:0031386
nucleosome, GO:0000786
nuclear nucleosome, GO:0000788
nuclear chromatin, GO:0000790
phosphatidylserine metabolism, GO:0006658
cell wall, GO:0005618
cell wall (sensu Fungi), GO:0009277
external encapsulating structure, GO:0030312
unfolded protein binding, GO:0051082
nucleoid, GO:0009295
mitochondrial nucleoid, GO:0042645
translational initiation, GO:0006413
translation initiation factor activity, GO:0003743
translation factor activity, nucleic acid binding, GO:0008135
translation regulator activity, GO:0045182
purine nucleoside triphosphate metabolism, GO:0009144
purine nucleoside triphosphate biosynthesis, GO:0009145
ribonucleoside triphosphate metabolism, GO:0009199
ribonucleoside triphosphate biosynthesis, GO:0009201
purine ribonucleoside triphosphate metabolism, GO:0009205
purine ribonucleoside triphosphate biosynthesis, GO:0009206
hydrogen transport, GO:0006818
proton transport, GO:0015992
nucleoside triphosphate biosynthesis, GO:0009142
nucleoside triphosphate metabolism, GO:0009141
purine ribonucleotide metabolism, GO:0009150
purine ribonucleotide biosynthesis, GO:0009152
ribonucleotide metabolism, GO:0009259
ribonucleotide biosynthesis, GO:0009260
purine nucleotide metabolism, GO:0006163
purine nucleotide biosynthesis, GO:0006164
monovalent inorganic cation transporter activity, GO:0015077
hydrogen ion transporter activity, GO:0015078
cation transporter activity, GO:0008324
ion transporter activity, GO:0015075
oxidative phosphorylation, GO:0006119
ATP synthesis coupled electron transport, GO:0042773
ATP synthesis coupled electron transport (sensu Eukaryota), GO:0042775
mitochondrial electron transport chain, GO:0005746
mitochondrial electron transport, cytochrome c to oxygen, GO:0006123
cytochrome-c oxidase activity, GO:0004129
respiratory chain complex IV (sensu Eukaryota), GO:0005751
heme-copper terminal oxidase activity, GO:0015002


1
2
3
4
5
6
7
8
9
10
Over-representation
5
-5
Under-representation
oxidoreductase activity, acting on heme group of donors, GO:0016675
oxidoreductase activity, acting on heme group of donors, oxygen as acceptor, GO:0016676
respiratory chain complex IV, GO:0045277
electron transport, GO:0006118
cellular respiration, GO:0045333
aerobic respiration, GO:0009060
mitochondrial membrane part, GO:0044455
phosphorylation, GO:0016310
phosphorus metabolism, GO:0006793
phosphate metabolism, GO:0006796
energy derivation by oxidation of organic compounds, GO:0015980
ribosomal small subunit assembly and maintenance, GO:0000028
hydrogen-transporting ATP synthase activity, rotational mechanism, GO:0046933
age-dependent general metabolic decline, GO:0007571
ribosome assembly, GO:0042255
ribosomal subunit assembly, GO:0042257
protein complex assembly, GO:0006461
ribosomal large subunit assembly and maintenance, GO:0000027
age-dependent response to reactive oxygen species, GO:0001315
age-dependent response to reactive oxygen species during chronological cell aging, GO:0001320
superoxide dismutase activity, GO:0004784
oxidoreductase activity, acting on superoxide radicals as acceptor, GO:0016721
condensed chromosome, GO:0000793
condensed nuclear chromosome, GO:0000794
condensed chromosome kinetochore, GO:0000777
condensed nuclear chromosome kinetochore, GO:0000778
spindle, GO:0005819
thiamin and derivative metabolism, GO:0042723
thiamin metabolism, GO:0006772
thiamin and derivative biosynthesis, GO:0042724
microtubule cytoskeleton, GO:0015630
chromosome, pericentric region, GO:0000775
spore wall assembly (sensu Fungi), GO:0030476
spore wall assembly, GO:0042244
meiosis, GO:0007126
meiotic cell cycle, GO:0051321
M phase of meiotic cell cycle, GO:0051327
condensed chromosome, pericentric region, GO:0000779
condensed nuclear chromosome, pericentric region, GO:0000780
chromosomal part, GO:0044427
proteasome complex (sensu Eukaryota), GO:0000502
cellular protein catabolism, GO:0044257
ubiquitin-dependent protein catabolism, GO:0006511
modification-dependent protein catabolism, GO:0019941
modification-dependent macromolecule catabolism, GO:0043632
proteolysis during cellular protein catabolism, GO:0051603
Golgi-associated vesicle, GO:0005798
proteasome core complex (sensu Eukaryota), GO:0005839
endopeptidase activity, GO:0004175
COPI vesicle coat, GO:0030126
COPI coated vesicle membrane, GO:0030663


1
2
3
4
5
6
7
8
9
10
Over-representation
5
Under-representation -5
proteasome core complex, beta-subunit complex (sensu Eukaryota), GO:0019774
proteasome core complex, alpha-subunit complex (sensu Eukaryota), GO:0019773
proteasome regulatory particle, base subcomplex (sensu Eukaryota), GO:0008540
ER to Golgi transport vesicle, GO:0030134
transport vesicle, GO:0030133
ER to Golgi transport vesicle membrane, GO:0012507
COPII vesicle coat, GO:0030127
Golgi-associated vesicle membrane, GO:0030660
membrane coat, GO:0030117
coated membrane, GO:0048475
cytoplasmic vesicle part, GO:0044433
vesicle coat, GO:0030120
vesicle membrane, GO:0012506
cytoplasmic vesicle membrane, GO:0030659
coated vesicle membrane, GO:0030662
Golgi apparatus part, GO:0044431
ER to Golgi vesicle-mediated transport, GO:0006888
coated vesicle, GO:0030135
cytoplasmic membrane-bound vesicle, GO:0016023
cytoplasmic vesicle, GO:0031410
vesicle, GO:0031982
membrane-bound vesicle, GO:0031988
actin cortical patch, GO:0030479
cortical cytoskeleton, GO:0030863
cortical actin cytoskeleton, GO:0030864
COPI-coated vesicle, GO:0030137
peptidase activity, GO:0008233
protein catabolism, GO:0030163
proteolysis, GO:0006508
negative regulation of cell organization and biogenesis, GO:0051129
tricarboxylic acid cycle intermediate metabolism, GO:0006100
ubiquitin cycle, GO:0006512
proteasome regulatory particle (sensu Eukaryota), GO:0005838
cell budding, GO:0007114
asexual reproduction, GO:0019954
actin filament depolymerization, GO:0030042
actin filament binding, GO:0051015
actin binding, GO:0003779
protein depolymerization, GO:0051261
ribosome export from nucleus, GO:0000054
Golgi vesicle transport, GO:0048193
Golgi apparatus, GO:0005794
mitochondrial matrix, GO:0005759
mitochondrial lumen, GO:0031980
nucleocytoplasmic transport, GO:0006913
nuclear transport, GO:0051169
cell division, GO:0051301
actin cytoskeleton, GO:0015629
cell cortex, GO:0005938
cell cortex part, GO:0044448
response to osmotic stress, GO:0006970
response to reactive oxygen species, GO:0000302

Supplementary Figure 1
Over- and underrepresented pathways in each of the expression clusters for wild-type yeast exposed to ethanol stress
This is the full list of pathways from Figure 3.3.

SUPPLEMENTARY TABLES

GO:0000027 ribosomal large subunit assembly and maintenance	GO:0015629 actin cytoskeleton
GO:0000082 G1/S transition of mitotic cell cycle	GO:0015630 microtubule cytoskeleton
GO:0000086 G2/M transition of mitotic cell cycle	GO:0015849 organic acid transport
GO:0000166 nucleotide binding	GO:0015934 large ribosomal subunit
GO:0000226 microtubule cytoskeleton organization and biogenesis	GO:0015935 small ribosomal subunit
GO:0000288 mRNA catabolism, deadenylylation-dependent decay	GO:0015980 energy derivation by oxidation of organic compounds
GO:0000293 ferric-chelate reductase activity	GO:0016023 cytoplasmic membrane-bound vesicle
GO:0000300 peripheral to membrane of membrane fraction	GO:0016209 antioxidant activity
GO:0000313 organellar ribosome	GO:0016282 eukaryotic 43S preinitiation complex
GO:0000502 proteasome complex (sensu Eukaryota)	GO:0016283 eukaryotic 48S initiation complex
GO:0000775 chromosome, pericentric region	GO:0016310 phosphorylation
GO:0000777 condensed chromosome kinetochore	GO:0016563 transcriptional activator activity
GO:0000778 condensed nuclear chromosome kinetochore	GO:0016620 oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor
GO:0000779 condensed chromosome, pericentric region	GO:0016675 oxidoreductase activity, acting on heme group of donors

GO:0000780 condensed nuclear chromosome, pericentric region	GO:0016676 oxidoreductase activity, acting on heme group of donors, oxygen as acceptor
GO:0000786 nucleosome	GO:0016758 transferase activity, transferring hexosyl groups
GO:0000788 nuclear nucleosome	GO:0016773 phosphotransferase activity, alcohol group as acceptor
GO:0000790 nuclear chromatin	GO:0016887 ATPase activity
GO:0000793 condensed chromosome	GO:0017038 protein import
GO:0000794 condensed nuclear chromosome	GO:0018456 aryl-alcohol dehydrogenase activity
GO:0003678 DNA helicase activity	GO:0019773 proteasome core complex, alpha-subunit complex (sensu Eukaryota)
GO:0003700 transcription factor activity	GO:0019774 proteasome core complex, beta-subunit complex (sensu Eukaryota)
GO:0004129 cytochrome-c oxidase activity	GO:0019783 small conjugating protein-specific protease activity
GO:0004175 endopeptidase activity	GO:0019941 modification-dependent protein catabolism
GO:0004386 helicase activity	GO:0019954 asexual reproduction
GO:0004672 protein kinase activity	GO:0030010 establishment of cell polarity
GO:0004812 aminoacyl-tRNA ligase activity	GO:0030029 actin filament-based process
GO:0005319 lipid transporter activity	GO:0030036 actin cytoskeleton organization and biogenesis
GO:0005618 cell wall	GO:0030117 membrane coat
GO:0005643 nuclear pore	GO:0030120 vesicle coat

GO:0005732 small nucleolar ribonucleoprotein complex	GO:0030126 COPI vesicle coat
GO:0005746 mitochondrial electron transport chain	GO:0030127 COPII vesicle coat
GO:0005751 respiratory chain complex IV (sensu Eukaryota)	GO:0030133 transport vesicle
GO:0005759 mitochondrial matrix	GO:0030134 ER to Golgi transport vesicle
GO:0005761 mitochondrial ribosome	GO:0030135 coated vesicle
GO:0005794 Golgi apparatus	GO:0030137 COPI-coated vesicle
GO:0005798 Golgi-associated vesicle	GO:0030163 protein catabolism
GO:0005819 spindle	GO:0030234 enzyme regulator activity
GO:0005830 cytosolic ribosome (sensu Eukaryota)	GO:0030312 external encapsulating structure
GO:0005839 proteasome core complex (sensu Eukaryota)	GO:0030427 site of polarized growth
GO:0005842 cytosolic large ribosomal subunit (sensu Eukaryota)	GO:0030467 establishment and/or maintenance of cell polarity (sensu Fungi)
GO:0005843 cytosolic small ribosomal subunit (sensu Eukaryota)	GO:0030468 establishment of cell polarity (sensu Fungi)
GO:0005933 bud	GO:0030476 spore wall assembly (sensu Fungi)
GO:0005938 cell cortex	GO:0030479 actin cortical patch
GO:0006092 main pathways of carbohydrate metabolism	GO:0030658 transport vesicle membrane
GO:0006096 glycolysis	GO:0030659 cytoplasmic vesicle membrane

GO:0006118 electron transport	GO:0030660 Golgi-associated vesicle membrane
GO:0006119 oxidative phosphorylation	GO:0030662 coated vesicle membrane
GO:0006123 mitochondrial electron transport, cytochrome c to oxygen	GO:0030663 COPI coated vesicle membrane
GO:0006272 leading strand elongation	GO:0030863 cortical cytoskeleton
GO:0006284 base-excision repair	GO:0030864 cortical actin cytoskeleton
GO:0006357 regulation of transcription from RNA polymerase II promoter	GO:0031124 mRNA 3'-end processing
GO:0006364 rRNA processing	GO:0031386 protein tag
GO:0006365 35S primary transcript processing	GO:0031410 cytoplasmic vesicle
GO:0006403 RNA localization	GO:0031980 mitochondrial lumen
GO:0006405 RNA export from nucleus	GO:0031982 vesicle
GO:0006413 translational initiation	GO:0031988 membrane-bound vesicle
GO:0006417 regulation of protein biosynthesis	GO:0035251 UDP-glucosyltransferase activity
GO:0006418 tRNA aminoacylation for protein translation	GO:0042026 protein refolding
GO:0006445 regulation of translation	GO:0042138 meiotic DNA double-strand break formation
GO:0006486 protein amino acid glycosylation	GO:0042175 nuclear envelope-endoplasmic reticulum network
GO:0006508 proteolysis	GO:0042244 spore wall assembly

GO:0006511 ubiquitin-dependent protein catabolism	GO:0042255 ribosome assembly
GO:0006512 ubiquitin cycle	GO:0042257 ribosomal subunit assembly
GO:0006520 amino acid metabolism	GO:0042274 ribosomal small subunit biogenesis
GO:0006551 leucine metabolism	GO:0042723 thiamin and derivative metabolism
GO:0006606 protein import into nucleus	GO:0042724 thiamin and derivative biosynthesis
GO:0006633 fatty acid biosynthesis	GO:0042773 ATP synthesis coupled electron transport
GO:0006665 sphingolipid metabolism	GO:0042775 ATP synthesis coupled electron transport (sensu Eukaryota)
GO:0006772 thiamin metabolism	GO:0043169 cation binding
GO:0006793 phosphorus metabolism	GO:0043487 regulation of RNA stability
GO:0006796 phosphate metabolism	GO:0043488 regulation of mRNA stability
GO:0006888 ER to Golgi vesicle-mediated transport	GO:0043632 modification-dependent macromolecule catabolism
GO:0006890 retrograde vesicle-mediated transport, Golgi to ER	GO:0044257 cellular protein catabolism
GO:0006897 endocytosis	GO:0044427 chromosomal part
GO:0006913 nucleocytoplasmic transport	GO:0044431 Golgi apparatus part
GO:0006970 response to osmotic stress	GO:0044433 cytoplasmic vesicle part
GO:0007114 cell budding	GO:0044437 vacuolar part

GO:0007126 meiosis	GO:0044445 cytosolic part
GO:0007163 establishment and/or maintenance of cell polarity	GO:0044448 cell cortex part
GO:0007571 age-dependent general metabolic decline	GO:0044455 mitochondrial membrane part
GO:0008026 ATP-dependent helicase activity	GO:0045277 respiratory chain complex IV
GO:0008233 peptidase activity	GO:0045333 cellular respiration
GO:0008324 cation transporter activity	GO:0045893 positive regulation of transcription, DNA-dependent
GO:0008519 ammonium transporter activity	GO:0046695 SLIK (SAGA-like) complex
GO:0008540 proteasome regulatory particle, base subcomplex (sensu Eukaryota)	GO:0046930 pore complex
GO:0008610 lipid biosynthesis	GO:0048193 Golgi vesicle transport
GO:0008639 small protein conjugating enzyme activity	GO:0048475 coated membrane
GO:0008652 amino acid biosynthesis	GO:0048519 negative regulation of biological process
GO:0009060 aerobic respiration	GO:0051119 sugar transporter activity
GO:0009071 serine family amino acid catabolism	GO:0051168 nuclear export
GO:0009100 glycoprotein metabolism	GO:0051169 nuclear transport
GO:0009277 cell wall (sensu Fungi)	GO:0051246 regulation of protein metabolism
GO:0012506 vesicle membrane	GO:0051252 regulation of RNA metabolism

GO:0012507 ER to Golgi transport vesicle membrane	GO:0051301 cell division
GO:0015002 heme-copper terminal oxidase activity	GO:0051321 meiotic cell cycle
GO:0015075 ion transporter activity	GO:0051327 M phase of meiotic cell cycle
GO:0015077 monovalent inorganic cation transporter activity	GO:0051603 proteolysis during cellular protein catabolism
GO:0015078 hydrogen ion transporter activity	

Supplementary Table 1
List of all gene ontology classes significantly upregulated or downregulated in response to threshold lethal ethanol stress in wild-type yeast

GO:0000027 ribosomal large subunit assembly and maintenance	GO:0008652 amino acid biosynthesis
GO:0000028 ribosomal small subunit assembly and maintenance	GO:0009066 aspartate family amino acid metabolism
GO:0000030 mannosyltransferase activity	GO:0009069 serine family amino acid metabolism
GO:0000096 sulfur amino acid metabolism	GO:0009070 serine family amino acid biosynthesis
GO:0000147 actin cortical patch assembly	GO:0009092 homoserine metabolism
GO:0000151 ubiquitin ligase complex	GO:0009100 glycoprotein metabolism
GO:0000221 hydrogen-transporting ATPase V1 domain	GO:0009101 glycoprotein biosynthesis
GO:0000502 proteasome complex (sensu Eukaryota)	GO:0009112 nucleobase metabolism
GO:0000793 condensed chromosome	GO:0009117 nucleotide metabolism
GO:0000794 condensed nuclear chromosome	GO:0009165 nucleotide biosynthesis
GO:0000819 sister chromatid segregation	GO:0009250 glucan biosynthesis
GO:0001405 presequence translocase-associated import motor	GO:0009295 nucleoid
GO:0003743 translation initiation factor activity	GO:0009309 amine biosynthesis
GO:0003777 microtubule motor activity	GO:0015077 monovalent inorganic cation transporter activity
GO:0004672 protein kinase activity	GO:0015078 hydrogen ion transporter activity
GO:0004722 protein serine/threonine phosphatase activity	GO:0015144 carbohydrate transporter activity

GO:0004812 aminoacyl-tRNA ligase activity	GO:0015662 ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism
GO:0004840 ubiquitin conjugating enzyme activity	GO:0015749 monosaccharide transport
GO:0004871 signal transducer activity	GO:0015934 large ribosomal subunit
GO:0005657 replication fork	GO:0015935 small ribosomal subunit
GO:0005732 small nucleolar ribonucleoprotein complex	GO:0016074 snoRNA metabolism
GO:0005736 DNA-directed RNA polymerase I complex	GO:0016282 eukaryotic 43S preinitiation complex
GO:0005744 mitochondrial inner membrane presequence translocase complex	GO:0016283 eukaryotic 48S initiation complex
GO:0005759 mitochondrial matrix	GO:0016471 hydrogen-translocating V-type ATPase complex
GO:0005830 cytosolic ribosome (sensu Eukaryota)	GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor
GO:0005838 proteasome regulatory particle (sensu Eukaryota)	GO:0016757 transferase activity, transferring glycosyl groups
GO:0005839 proteasome core complex (sensu Eukaryota)	GO:0016758 transferase activity, transferring hexosyl groups
GO:0005842 cytosolic large ribosomal subunit (sensu Eukaryota)	GO:0016763 transferase activity, transferring pentosyl groups
GO:0005843 cytosolic small ribosomal subunit (sensu Eukaryota)	GO:0016820 hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances
GO:0005871 kinesin complex	GO:0016866 intramolecular transferase activity

GO:0006007 glucose catabolism	GO:0016875 ligase activity, forming carbon-oxygen bonds
GO:0006096 glycolysis	GO:0016876 ligase activity, forming aminoacyl-tRNA and related compounds
GO:0006275 regulation of DNA replication	GO:0019320 hexose catabolism
GO:0006281 DNA repair	GO:0030126 COPI vesicle coat
GO:0006357 regulation of transcription from RNA polymerase II promoter	GO:0030150 protein import into mitochondrial matrix
GO:0006364 rRNA processing	GO:0030476 spore wall assembly (sensu Fungi)
GO:0006365 35S primary transcript processing	GO:0030663 COPI coated vesicle membrane
GO:0006399 tRNA metabolism	GO:0030684 preribosome
GO:0006413 translational initiation	GO:0030685 nucleolar preribosome
GO:0006417 regulation of protein biosynthesis	GO:0031980 mitochondrial lumen
GO:0006418 tRNA aminoacylation for protein translation	GO:0042244 spore wall assembly
GO:0006445 regulation of translation	GO:0042255 ribosome assembly
GO:0006450 regulation of translational fidelity	GO:0042257 ribosomal subunit assembly
GO:0006457 protein folding	GO:0042626 ATPase activity, coupled to transmembrane movement of substances
GO:0006458 'de novo' protein folding	GO:0042645 mitochondrial nucleoid
GO:0006461 protein complex assembly	GO:0043038 amino acid activation

GO:0006486 protein amino acid glycosylation	GO:0043039 tRNA aminoacylation
GO:0006493 protein amino acid O-linked glycosylation	GO:0043413 biopolymer glycosylation
GO:0006519 amino acid and derivative metabolism	GO:0043492 ATPase activity, coupled to movement of substances
GO:0006520 amino acid metabolism	GO:0044271 nitrogen compound biosynthesis
GO:0006566 threonine metabolism	GO:0044445 cytosolic part
GO:0006626 protein targeting to mitochondrion	GO:0044452 nucleolar part
GO:0006725 aromatic compound metabolism	GO:0044455 mitochondrial membrane part
GO:0007009 plasma membrane organization and biogenesis	GO:0045182 translation regulator activity
GO:0007126 meiosis	GO:0045332 phospholipid translocation
GO:0007154 cell communication	GO:0046164 alcohol catabolism
GO:0007165 signal transduction	GO:0046365 monosaccharide catabolism
GO:0008028 monocarboxylic acid transporter activity	GO:0046961 hydrogen-transporting ATPase activity, rotational mechanism
GO:0008135 translation factor activity, nucleic acid binding	GO:0051083 cotranslational protein folding
GO:0008540 proteasome regulatory particle, base subcomplex (sensu Eukaryota)	GO:0051246 regulation of protein metabolism
GO:0008639 small protein conjugating enzyme activity	GO:0051321 meiotic cell cycle
GO:0008643 carbohydrate transport	GO:0051327 M phase of meiotic cell cycle

GO:0008645 hexose transport	

Supplementary Table 2
List of all gene ontology classes significantly increased or decreased in strain JY304 compared the wild-type parental strain before the onset of threshold lethal ethanol stress

APPENDIX B: OTHER COLLABORATIVE WORK

In addition to my thesis project, I also worked closely with a former postdoctoral researcher in the lab, Peter Freddolino, who is now an assistant professor at the University of Michigan. The publication that resulted from this work, titled “Cellular Adaptation Through Fitness-Directed Transcriptional Tuning,” is attached directly after this page.

eLIFE elifesciences.org

RESEARCH ARTICLE

Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry

Peter L Freddolino[1,2†], Jamie Yang[1,2], Amir Momen-Roknabadi[1,2], Saeed Tavazoie[1,2*]

[1]Department of Systems Biology, Columbia University, New York City, United States; [2]Department of Biochemistry and Molecular Biophysics, Columbia University, New York City, United States

Abstract Cells adapt to familiar changes in their environment by activating predefined regulatory programs that establish adaptive gene expression states. These hard-wired pathways, however, may be inadequate for adaptation to environments never encountered before. Here, we reveal evidence for an alternative mode of gene regulation that enables adaptation to adverse conditions without relying on external sensory information or genetically predetermined *cis*-regulation. Instead, individual genes achieve optimal expression levels through a stochastic search for improved fitness. By focusing on improving the overall health of the cell, the proposed stochastic tuning mechanism discovers global gene expression states that are fundamentally new and yet optimized for novel environments. We provide experimental evidence for stochastic tuning in the adaptation of *Saccharomyces cerevisiae* to laboratory-engineered environments that are foreign to its native gene-regulatory network. Stochastic tuning operates locally at individual gene promoters, and its efficacy is modulated by perturbations to chromatin modification machinery.
DOI: https://doi.org/10.7554/eLife.31867.001

***For correspondence:** st2744@columbia.edu

Present address: †Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, United States

Competing interests: The authors declare that no competing interests exist.

Funding: See page 28

Received: 09 September 2017
Accepted: 04 April 2018
Published: 05 April 2018

Reviewing editor: Naama Barkai, Weizmann Institute of Science, Israel

Introduction

The capacity to adapt to changes in the external environment is a defining feature of living systems. Cells can rapidly adapt to familiar changes that are commonly encountered in their native habitat by sensing the parameters of the environment and engaging dedicated regulatory networks that have evolved to establish adaptive gene expression states (***Jacob and Monod, 1961***; ***Thieffry et al., 1998***). However, dedicated sensory, signaling, and regulatory networks become inadequate, or even detrimental, when cells are exposed to unfamiliar environments that are foreign to their evolutionary history (***Tagkopoulos et al., 2008***). In principle, at least one gene expression state that maximizes the health/fitness of the cell always exists, despite the inability of the native regulatory network to establish such a state. This is true because under any conceivable environment, the activities of some genes are beneficial, whereas those of others are futile or even actively detrimental (***Jacob and Monod, 1961***; ***Tagkopoulos et al., 2008***; ***Hottes et al., 2013***). In fact, if the initial fitness defect is not lethal, a population of cells may slowly adapt to an unfamiliar environment through the accumulation of genetic mutations that rewire regulatory networks, thereby achieving more optimal gene expression states (***Tagkopoulos et al., 2008***; ***Applebee et al., 2008***; ***Philippe et al., 2007***; ***Goodarzi et al., 2010***; ***Tenaillon et al., 2012***; ***Rodríguez-Verdugo et al., 2016***; ***Blount et al., 2012***; ***Van Hofwegen et al., 2016***; ***Damkiær et al., 2013***).

eLife digest To survive, cells have to adapt to changes in their environment. Organisms can do so by constantly modifying the expression of their genes. For example, bacteria exposed to high temperatures turn on heat-shock genes to help them cope.

Responses to familiar environmental changes take place thanks to specific, hard-wired molecular pathways. These transmit external signals to transcription factors, proteins that can bind DNA near a gene to regulate its expression. Yet, such established responses may not exist for stressful conditions that cells have never encountered during their evolutionary history. In this case, how can organisms adjust which genes to express, and at what levels?

Here, Freddolino et al. theorize that, in a new environment, individual genes can randomly increase or decrease their level of expression. If a change ends up being good for the survival of the cell, it is further reinforced. This 'stochastic tuning' would allow organisms to find the optimal levels of gene expression without using genetically predetermined pathways that involve transcription factors.

Mathematical simulations suggest that this mechanism can improve the growth and survival of a cell in a new environment. Diverse experiments demonstrate that a phenomenon consistent with stochastic tuning occurs in yeasts. The organisms are genetically modified so that their transcription factors can no longer activate *URA3*, a gene required to grow in conditions lacking a chemical called uracil. Yet, these altered yeast cells still manage to boost their *URA3* expression in a uracil-free environment.

Stochastic tuning could thus work alongside other types of conventional gene regulation to help cells adapt to new and challenging living conditions. For instance, this may be how cancerous cells survive and thrive when facing chemotherapy drugs.

DOI: https://doi.org/10.7554/eLife.31867.002

Results

Adaptation through fitness-driven stochastic optimization of gene expression

In this work we speculate whether cells have evolved alternative strategies for finding adaptive gene expression states, on more physiological timescales, without relying on their hard-coded sensory and regulatory systems. Since the perception of the external world may be of limited value under unfamiliar conditions, perhaps a more effective strategy would be to focus on maximizing the internal health of the cell—without regard to the specific parameters of the outside world. This would be a challenging strategy, as every gene in the genome would need to independently reach the expression level that maximizes the overall health of the cell, and these expression levels could vary significantly from condition to condition. In particular, we asked whether individual genes could, in principle, carry out a search process equivalent to gradient descent (***Cauchy, 1847***), where the health consequence of stochastic alterations in gene expression could gradually tune the expression of individual genes towards a level that is optimal for internal health. We reasoned that such an optimization process would require the existence of: (1) a source of stochastic transitions in gene expression; (2) the ability of local chromatin to maintain a record of recent changes in transcription; and (3) a central metabolic hub that integrates diverse parameters of intracellular health and continuously broadcasts whether the overall health of the cell is improving or deteriorating. In fact, we find that the foundations for meeting these requirements are already present in eukaryotic cells: (1) The expression of many genes is dominated by noisy bursts of transcription—a widespread phenomenon of largely unknown functional significance (***Sanchez and Golding, 2013***; ***Raj and van Oudenaarden, 2008***; ***Blake et al., 2006***; ***Raser and O'Shea, 2005***; ***Elowitz et al., 2002***); (2) Co-transcriptional histone modification can modify eukaryotic chromatin in promoters and gene bodies, establishing a short-term memory of recent transcriptional events (***Li et al., 2007***; ***Rando and Winston, 2012***); and (3) Global integrators of cell health have evolved in eukaryotes. A classic example is the mTOR pathway, which integrates a vast array of intracellular parameters reflecting nutrient availability, energy,

and the presence of diverse stresses (***Conrad et al., 2014***; ***González and Hall, 2017***; ***Albert and Hall, 2015***; ***Saxton and Sabatini, 2017***).

With the necessary components for gradient-based optimization of gene expression in place (***Figure 1A***), the promoter of each gene would be able to conduct a simple search process that culminates in finding the expression level that maximizes the overall health of the cell: if global fitness/health is increasing *and* there was a previous increase in transcriptional output (representing larger or more frequent transcriptional bursts), the promoter further increases its transcriptional activity (***Figure 1B***). If fitness is decreasing *and* there was a previous increase in transcriptional activity, the promoter decreases its transcriptional output. Transcriptional output is altered in the opposite direction in the event that there was a previous decrease in transcriptional output. For each gene, this tuning process can be expressed as: $\Delta E_t = k \cdot sgn(\Delta F_t \cdot \Delta E_{t-1}) + \eta$ (see ***Figures 1*** and ***2A***); here, *E* denotes the vector of gene-level transcription rates, *F* the current fitness/health of the cell, *k* is a proportionality constant, η a noise term, and *sgn* is a function yielding -1 if its argument is negative, 0 if its argument is zero, and +1 if its argument is positive. One can easily see how the process described here can tune the optimal expression of a single gene. What is remarkable, however, is the ability of this hypothetical stochastic tuning process to find near-optimal gene-expression states for a system with thousands of genes. As can be seen in the simulations presented in ***Figure 2***, this is achieved through a fitness-directed stochastic search culminating in individual genes reaching specific gene expression levels that maximize the health/fitness of the cell. Such a stochastic tuning

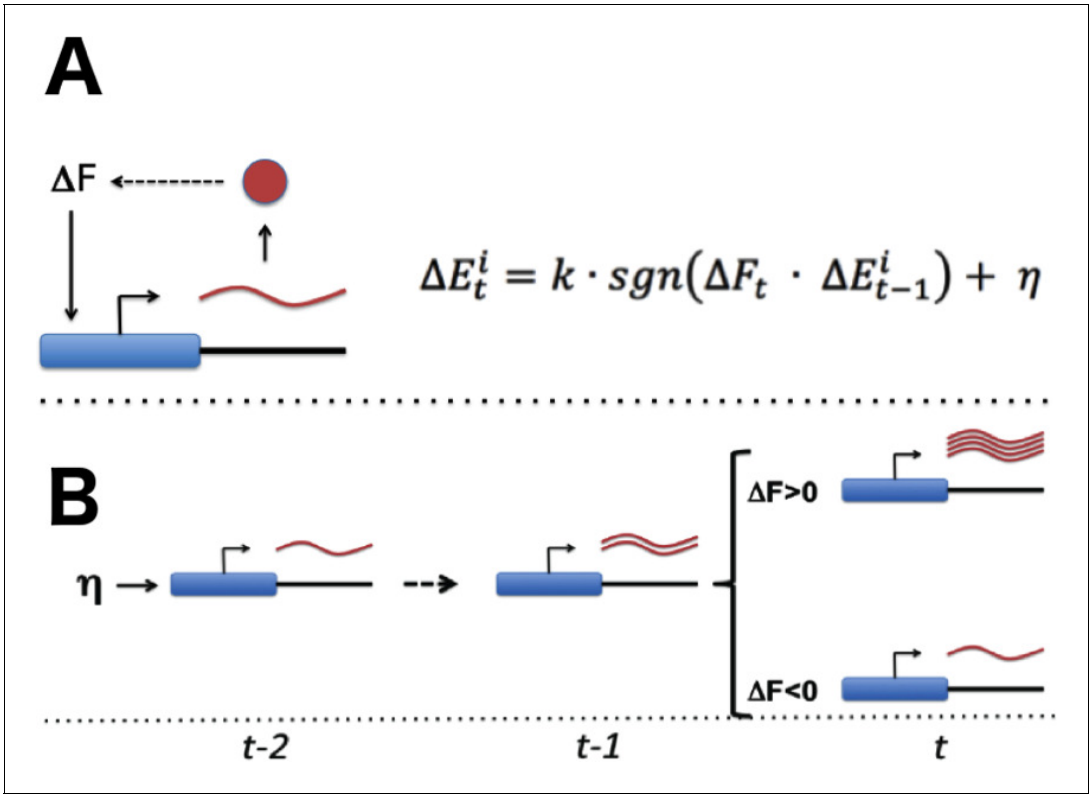


Figure 1. Stochastic tuning of gene expression by fitness optimization at gene promoters. (**A**) Each gene contains a noisy expression apparatus with noise amplitude η that allows exploration of a range of transcriptional activities. Each transcription apparatus also maintains a record of its previous change in transcriptional activity (ΔE_{t-1}). The change in transcriptional activity has the potential to contribute to a change in global health (ΔF_t) through the downstream effect of the gene product's activity (likely through a multi-step pathway; for example, the biosynthesis of a metabolite that is limiting for growth). A global metabolic integrator can transduce this change in health/fitness to every gene's expression apparatus. At any point in time, the expression apparatus executes a change in transcriptional activity (ΔE_t) proportional (k) to the sign (*sgn*) of the product of ΔE_{t-1} and ΔF_t plus noise (η). (**B**) A simple example of this can be seen for a gene that experiences a random burst in transcriptional activity. If this leads to an increase in fitness the expression apparatus further increases transcriptional activity. Conversely, if there is a decrease in fitness, the expression apparatus decreases transcriptional activity.

DOI: https://doi.org/10.7554/eLife.31867.003

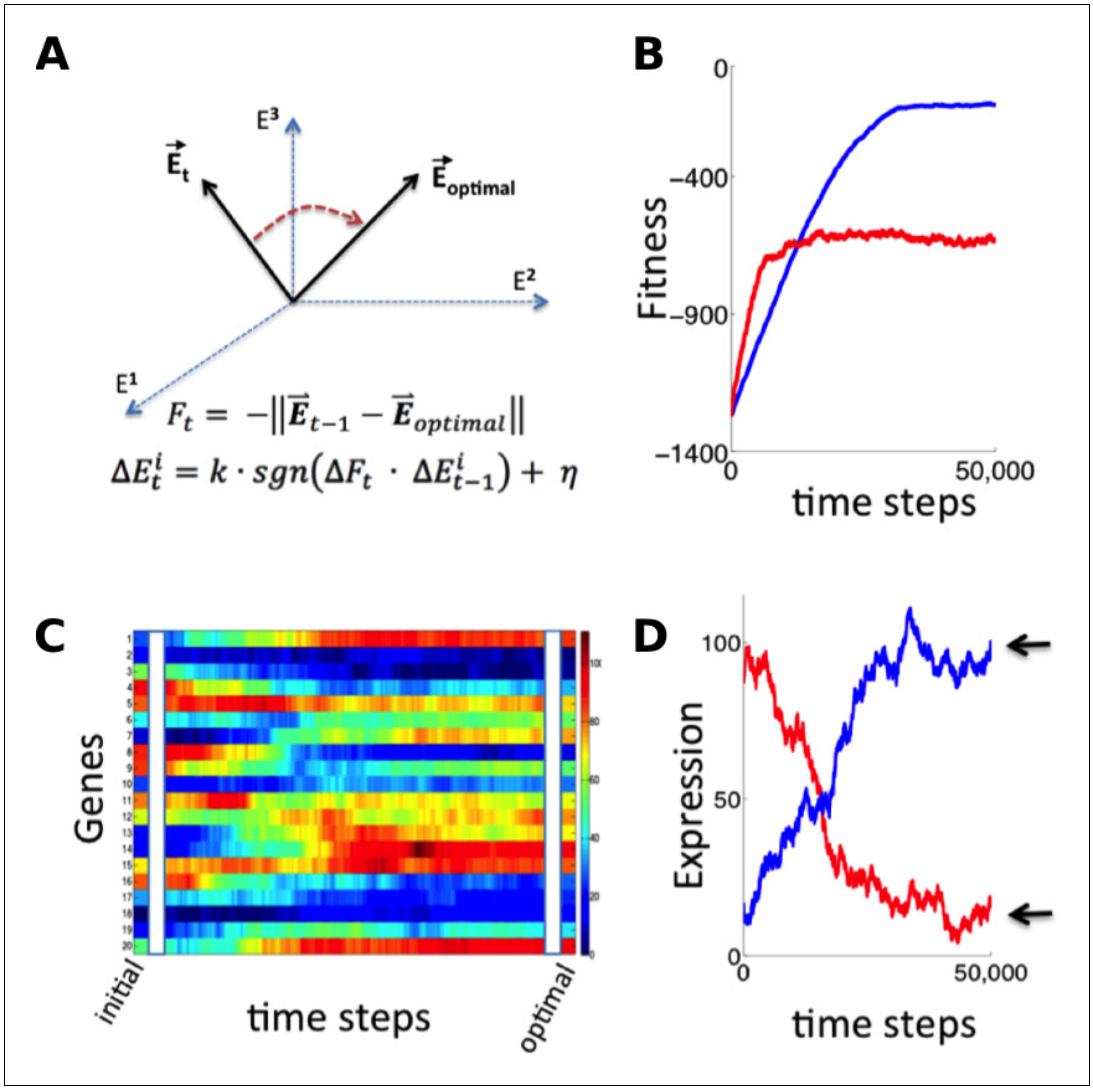


Figure 2. Simulation of fitness-directed stochastic tuning for a thousand-gene system. (**A**) Quantitative framework describing stochastic tuning. The transcriptional activity state of the genome is represented by the vector E, here schematically represented for a three-gene system. In any environment, there is an optimal transcriptional state vector ($E_{optimal}$) that yields maximum fitness. At any time (t), a cell with transcriptional activity state E_t has global health/fitness (F_t) defined as the negative of the Euclidean distance between the immediately preceding transcriptional activity state E_{t-1} and $E_{optimal}$. Each gene promoter (i) executes a change in transcriptional activity ΔE^i_t which has two components: (1) a step with magnitude of *k* and sign (*sgn*) matching that of the product of the global change in fitness (ΔF_t) experienced at time *t* and the preceding change in transcriptional activity ΔE^i_{t-1}, and (2) a noise component with a magnitude of η and a random sign (+/-). (**B**) The stochastic tuning process moves the transcriptional activity state towards the optimum, resulting in increasing health/fitness over time. Simulated trajectories are shown for a 1,000-gene system with $k = 0.1$, $\eta = 0.1$ (blue); $k = 0.5$, $\eta = 0.5$ (red). (**C**) The time evolution of the transcriptional activity state vector as a system containing 1000 genes converges to optimal transcriptional activities through stochastic tuning. The temporal profiles of 20 representative genes are shown, starting from randomly assigned initial activities, and gradually converging to activities that are near optimal for fitness (using parameters corresponding to the blue curve in panel **B**). (**D**) Trajectories of two representative genes are shown for the same simulation as in panel C). Transcriptional activities start at randomly assigned initial values and gradually converge to near the optimum (arrows).

DOI: https://doi.org/10.7554/eLife.31867.004

mechanism would be highly valuable to free-living organisms, enabling them to optimize their global gene expression patterns to match the specific requirements of any environment in which their dedicated sensory and regulatory networks are inadequate or sub-optimal.

Fitness-directed tuning of gene expression in yeast

Informed by the simulations above, we sought to test for evidence of stochastic tuning in the eukaryotic model organism *Saccharomyces cerevisiae*. We engineered conditions in which the expression of a single gene was required for growth, but for which no regulatory input existed to drive appropriate expression levels. This was achieved by using a yeast strain (BY4743) that lacks the URA3 gene, which is essential when cells are grown in the absence of uracil. We placed a chromosomally integrated copy of URA3 at a different locus under the control of a weak synthetic promoter, consisting primarily of a pseudorandom sequence. All recognizable binding sites for native transcription factors were removed from the generated promoter sequence (see Materials and Methods and ***Supplementary file 1*** for details), in an attempt to decouple it from any existing sensory and regulatory input. We henceforth refer to this synthetic promoter sequence as synprom (see ***Supplementary file 1*** for sequence). In the experiments described below, URA3 is typically tagged with a fluorescent fusion, either mRuby (***Kredel et al., 2009***) or a superfolder GFP (***Pédelacq et al., 2006***), and a copy of a mouse DHFR gene coupled to a different fluorescent protein is inserted at the same location on the sister chromosome to act as an internal control. A schematic of the insertion constructs is shown in ***Figure 3A***. We also added the URA3 competitive antagonist 6-azauracil (6AU) to the media to control the threshold level of URA3 production required for growth. The growth condition, SC+glu-ura media, containing *x* µg/ml of 6AU, will henceforth be referred to as ura-/6AU*x*.

Even with the challenging and specific experimental layout described here, with growth highly dependent on URA3 expression, we expect that stochastic tuning might contribute to fitness through mechanisms acting in cis at the promoter driving URA3, those acting in trans through modulation of factors that (despite our best efforts) weakly affect the promoter driving URA3, and through tuning of unrelated pathways that benefit survival and growth in the –URA condition. Nevertheless, URA3 expression itself will clearly be the key driver of growth since it is the critical bottleneck for nucleotide biosynthesis in the absence of uracil supplementation.

To look for evidence of fitness-directed stochastic tuning, we tracked the colony formation of cells containing synprom-driven URA3 after plating on ura-/6AU15 plates. Lacking sufficient URA3 expression to overcome high 6AU levels, these non-growing cells would be expected to succumb to starvation and die. Remarkably, however, after prolonged incubation we observed apparently stochastic transitions to rapid growth, leading to the formation of macroscopic colonies over time (***Figure 3B***). We eventually observed colony formation by roughly one cell in 10^3, a rate too high to be driven by mutation-driven adaptation in the absence of growth.

Stochastic tuning of other synthetic and natural promoters

The synthetic promoter referred to as 'synprom' throughout the text is the combination of a pseudorandom sequence with a small natural promoter-proximal region taken from the SAM3 gene, with both stripped of all recognizable matches to known transcription factor binding sites (see Materials and Methods for details). We also tested all combinations of five other synthetic promoter sequences and one other promoter proximal region, enumerated in ***Supplementary file 2***. As shown in ***Figure 3—figure supplement 1***, four of the six synthetic promoters support stochastic tuning, and the ability of synprom5 (the purely artificial component of the synprom referred to in the remainder of the text; see ***Supplementary file 2*** for all synthetic promoter sequences) to undergo tuning remains even with a different promoter proximal region. These findings highlight the universality of the observed tuning phenomenon and minimize the possibility that our observations actually arise due to the presence of some residual sequence-specific transcription factor binding site present in synprom.

As shown in ***Figure 3C***, we also observed similar tuning behavior for two high-noise natural promoters, P_{HSP12} and P_{RGI1} (***Tirosh et al., 2009***; ***Tirosh et al., 2006***), indicating that stochastic tuning can function even when superimposed on naturally evolved regulatory sites. Across all promoters

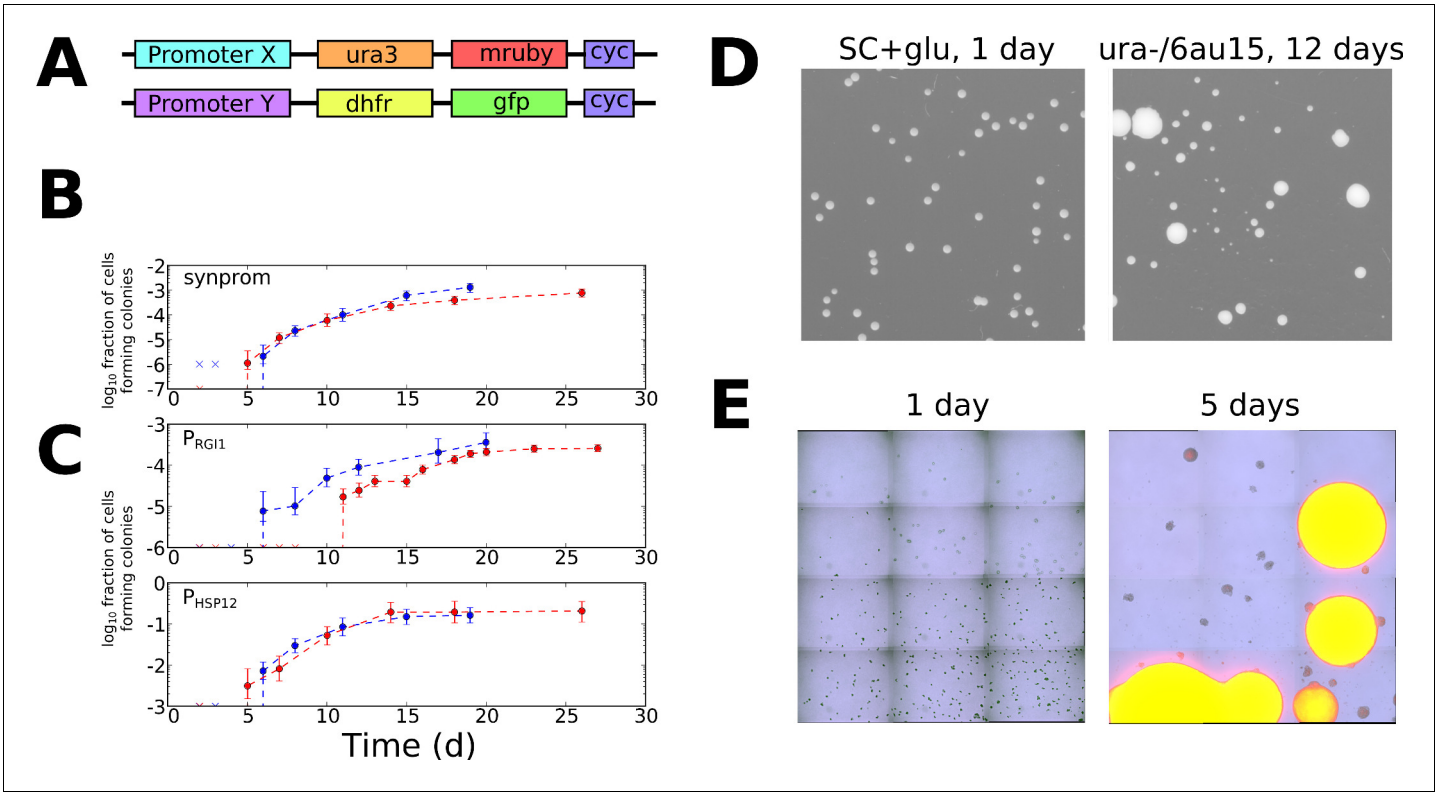


Figure 3. Stochastic tuning of yeast cells under uracil starvation. (**A**) Schematic of the constructs used in this study. All strains are diploid, containing similar insertions at the LEU2 locus of both copies of chromosome III. X is either a synthetic promoter (synprom) or a natural promoter (P_{RGI1} or P_{HSP12}) unless otherwise noted, and Y is either the same promoter as X or is the strong constitutive promoter P_{ADH1}. 'cyc' indicates the well-characterized CYC1 transcriptional terminator (***Russo and Sherman, 1989***). (**B**) Stochastic colony formation on ura-/6AU15 plates for cells containing URA3-mRuby under control of synprom and DHFR-GFP under control of P_{ADH1}. Error bars show central 95% credible intervals; colors show biological replicates performed on different days. 'x' marks are shown at the bottom of the axis for days where zero visible colonies were present at all plated dilutions. Cells plated on SC+glu uniformly form visible colonies within 1–2 days. (**C**) As in panel B, but with URA3-mRuby controlled by P_{RGI1} or P_{HSP12} as indicated. (**D**) Images of colony growth on SC+glu and ura-/6AU15 plates taken at the specified number of days after plating (1 day for SC+glu, 12 days for ura-/6AU15). Growth of colonies is nearly uniform on SC+glu plates but shows non-uniform stochastic emergence on ura-/6AU15. *N.b.* the plated dilutions for the two plate types are not the same. URA3 expression for the experiment shown is controlled by P_{HSP12}, but similar behavior was observed for all promoters discussed here. (**E**) Early colony formation on ura-/6AU15 plates imaged by superimposed differential interference contrast and fluorescence microscopy. Cells contain P_{HSP12}-URA3-mRuby/P_{ADH1}-DHFR-GFP. Left panel: One day after plating. By this timepoint small, macroscopic colonies would have formed on SC+glu plates, but instead cells remain in microcolonies having undergone no more than three doublings. Right panel: Same plate as left, five days after plating. While most cells have not grown since the one-day timepoint, other cells having undergone successful tuning instead form larger colonies with URA3 expression sustained throughout them.

DOI: https://doi.org/10.7554/eLife.31867.005

The following figure supplement is available for figure 3:

Figure supplement 1. Stochastic colony formation rates for cells with URA3 driven by a variety of synthetic promoters.

DOI: https://doi.org/10.7554/eLife.31867.006

(natural and synthetic) tested here, the observed tuning rates, relative to the number of viable plated cells, varied from 1 in 10^1 (P_{HSP12}) to 1 in 10^5 (synprom5-arf1).

The apparently stochastic nature of colony formation in our experiments is reflected both in the steady emergence of colonies over the course of days or weeks (***Figure 3B–C*** and ***Figure 3—figure supplement 1***), and in the wide variance of colony sizes observed on ura-/6AU15 plates (***Figure 3D***). Microscopy revealed that cells remain quiescent for days before transitioning to URA3 expression and rapid growth, with a transition rate dependent on the choice of promoter (***Figure 3E***). Furthermore, the change that enables growth under the ura-/6AU15 condition must be passed from mother to daughter cells, as colonies expand from a few points of initiation instead of showing random division of cells throughout the microscopic field over time. While the presence of some deterministic

process, yielding colony formation over the observed timescales (dependent on the initial state of each cell), cannot be ruled out, a far simpler explanation for the observed phenomenon of a long lag followed by appearance of colonies over a wide range of times is that each cell independently undergoes a random process that can eventually lead to growth. We confirmed that the appearance of colonies is not simply due to aging of the plates; 6AU-containing plates which were pre-incubated for a week or longer prior to plating of cells showed no change in colony formation rates (data not shown).

Fitness-directed tuning operates independently of conventional regulatory input and is transcriptionally driven

To provide further insights into the regulatory changes occurring during the onset of cell growth, we performed flow cytometry time courses on cells challenged by, and subsequently growing in, liquid ura-/6AU5 media, using cells with synprom-driven URA3-mRuby, and with a DHFR-GFP fusion driven by either the constitutive ADH1 promoter (***Figure 4A–B***) or synprom (***Figure 4C–D***) itself. The use of P_{ADH1} to drive the second reporter allows us to control for extrinsic noise and global changes in gene expression, whereas coupling synprom to the non-beneficial DHFR-GFP fusion allows us to test whether the observed stochastic tuning is driven by any *trans*-acting input from some existing

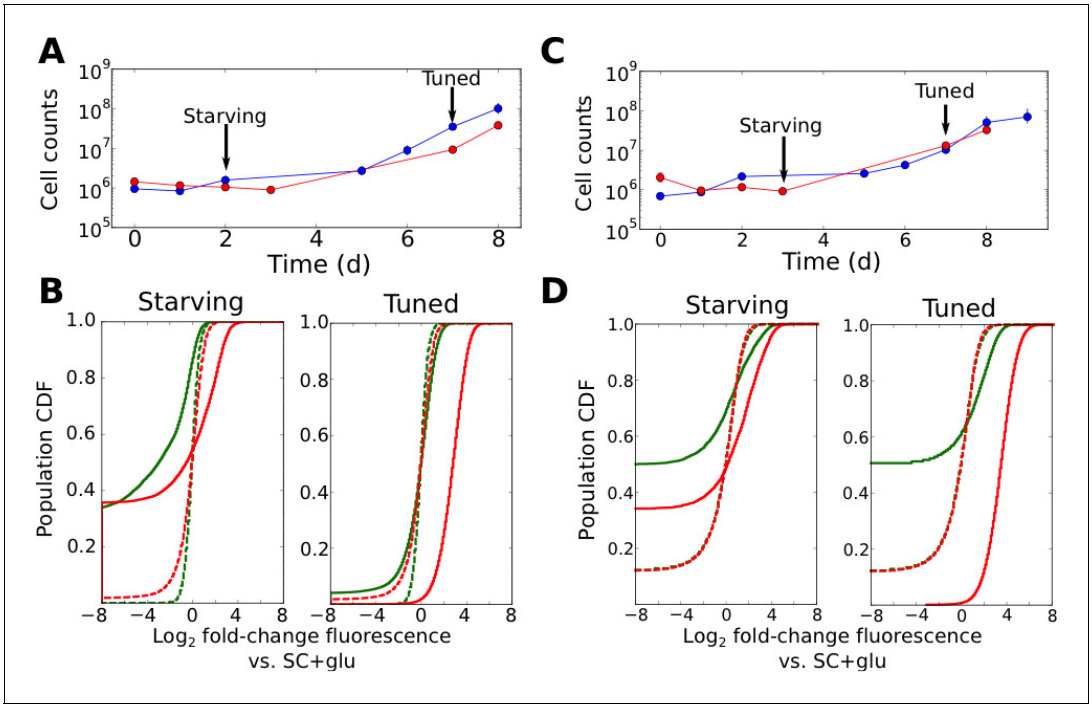


Figure 4. Tuning is both promoter- and allele-specific. (**A**) Cell counts for synprom-URA3-mRuby/P_{ADH1}-DHFR-GFP cells in liquid ura-/6AU5 media. Colors correspond to different biological replicates started on different days. Arrows indicate two timepoints from each strain for which fluorescence cumulative distribution functions (CDFs) are shown below. Error bars for cell counts show central 95% credible intervals. (**B**) Flow cytometry cumulative distributions of fluorescence levels for URA3-mRuby and DHFR-GFP during uracil starvation. In each CDF a given timepoint (solid line) is compared to the distribution present for cells in logarithmic growth in SC+glu (rich) media (dashed lines). The values shown are $\log_2$ ratios to the median value of cells growing exponentially in SC+glu. GFP signals are shown in green and mRuby signals in red. (**C**) Analogous to A, but we consider cells where synprom drives both URA3-mRuby and DHFR-GFP. (**D**) Analogous to B, but for cells with synprom driving both URA3-mRuby and DHFR-GFP.

DOI: https://doi.org/10.7554/eLife.31867.007

The following figure supplements are available for figure 4:

Figure supplement 1. Promoter-specific stochastic tuning of URA3 expression by native promoters in *S. cerevisiae*.

DOI: https://doi.org/10.7554/eLife.31867.008

Figure supplement 2. Local tuning of URA3 expression.

DOI: https://doi.org/10.7554/eLife.31867.009

regulatory network or whether it is truly specific to the allele needed for growth, as required by our proposed tuning model.

Several patterns in the growth curves and flow cytometry data are immediately apparent. First, as with the agar-based growth discussed above, cells show a lag of at least 72 hr with undetectable growth, followed by the onset of steady growth until saturation. In the case of URA3-mRuby driven by synprom and DHFR-GFP by the constitutive promoter P_{ADH1}, URA3-mRuby fluorescence increases substantially in tandem with the onset of cell growth, and expression subsequently remains high until saturation; in contrast, DHFR-GFP signals do not even recover to their initial levels (***Figure 4B***; compare dashed and solid line distributions). This demonstrates that the URA3 induction resulting in growth is promoter-specific and does not simply reflect a general increase in protein expression. We observed qualitatively equivalent behavior when URA3 was driven by P_{RGI1} or P_{HSP12} (***Figure 4—figure supplement 1***). Even more strikingly, for cells with synprom driving both fluorescent fusions, we observed a specific enhancement of URA3-mRuby expression over that of DHFR-GFP (***Figure 4D***), showing that the transition to high URA3 expression is not only promoter-specific but allele-specific, and thus must be driven at least partly by changes occurring in cis at the specific locus whose expression is required for growth. As an additional test, we performed quantitative RT-PCR experiments to measure the ratio of URA3 and DHFR mRNA expression in tuned cells either in liquid ura-/6AU5 media or on ura-/6AU15 plates (see ***Figure 4—figure supplement 2***). In both cases, we observed a substantial increase in the URA3:DHFR ratio in the tuned cells, indicating that the observed tuning occurs at least partly through a local *cis*-acting process at the locus required for growth (although we cannot rule out additional changes in other promoters that also contribute to survival and growth, which may account for the observed heterogeneity in expression levels between replicates). Consistent with our proposed tuning model, the allele-specific nature of the transcriptional induction supports a key role for a local tuning process that is independent of dedicated sensory and regulatory input.

Varying the threshold level of URA3 required for growth shifts tuning from stochastic to deterministic

The presence of the competitive URA3 inhibitor 6-azauracil allows us to vary the threshold level of URA3 required for growth. Thus, it is instructive to consider how the concentration of 6AU may alter stochastic tuning behavior, both in the context of the computational model described above and in the actual behavior of the system. We made two crucial modifications to the numerical model employed in ***Figure 2*** to mimic our experimental setup. First, rather than having the entire gene expression profile begin far from the optimal point, we begin with all genes but one (representing URA3) at their optimal values, reflecting the fact that aside from the artificial stress of lacking appropriate URA3 regulation, the cells' native regulatory network can provide an appropriate response to ura-/6AU media. Second, we note that due to the presence of the competitive inhibitor 6AU, the URA3 in the cell will not even be able to contribute meaningfully to nucleotide biosynthesis (and thus impact the cell's health/fitness) until it passes a threshold level. Thus, the tuning term (***Figure 2A***) is not applied to the gene representing URA3 until after the concentration of URA3 passes a threshold. Aside from the modifications noted above, we model tuning in the ura-/6AU environment as we did for the general case in ***Figure 2***, and in particular, the fitness effects of changing URA3 expression must compete with noisy gene expression from the other 999 genes in the model gene expression profile to impact the direction of tuning.

The resulting URA3 expression profiles during simulated tuning in the presence of low or high concentrations of 6AU are shown in ***Figure 5A***. In the low 6AU case, the tuning mechanism pushes URA3 expression almost deterministically to its optimal (high) value, whereas in the presence of high 6AU, the URA3 expression level undergoes a random walk until expression becomes high enough to allow the tuning mechanism to 'sense' the gradient and drive the cells into a URA3+ state. The effects on tuning rates of varying the 6AU concentration are plotted in ***Figure 5B***, where we observe that increasing 6AU concentrations both slow tuning and dramatically increase the variance in the amount of time required for each individual cell to reach a URA+ state. This is precisely the behavior observed experimentally with high 6AU concentrations (***Figure 3***). On the other hand, tuning in our experimental system switched from slow and stochastic to rapid and deterministic in the presence of low 6AU concentrations, with observable tuning occurring over the course of a few hours (***Figure 5C***). Importantly, the tuning process is confined to the URA3-mRuby allele, despite the fact

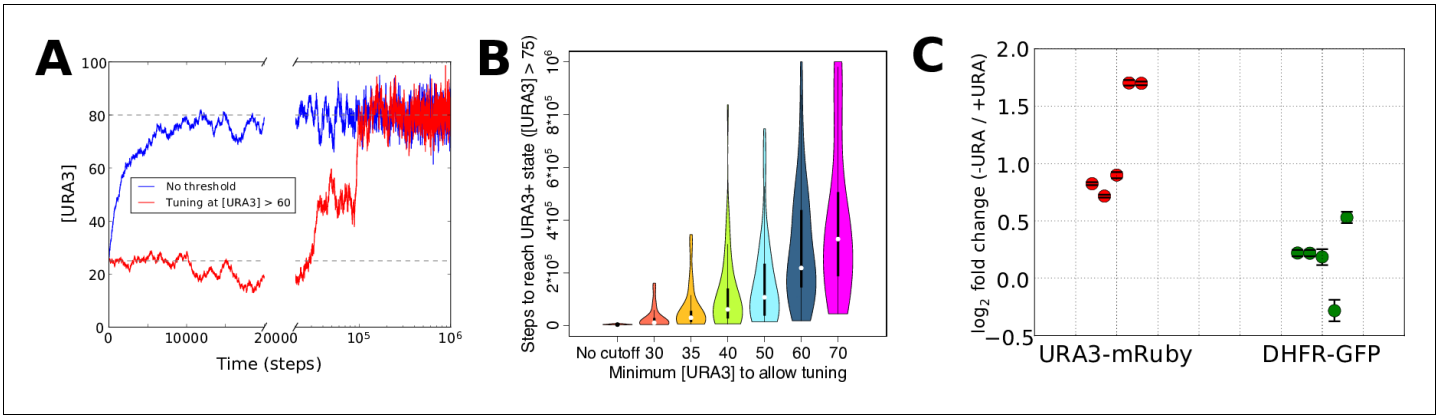


Figure 5. Numerical modeling and experimental validation of changes in tuning behavior as a function of 6AU concentration. We simulated the gene expression dynamics of cells containing URA3 under the control of a non-native promoter, when exposed to uracil-depletion stress with varying concentrations of the URA3 inhibitor 6AU. The model employed is equivalent to that in ***Figure 2A***, with $k = 0.1$, $\eta = 0.1$, and the target expression profile equal to that for the case shown in ***Figure 2B*** except for the case of the gene corresponding to URA3, whose optimal value was set to 80. (**A**) Typical trajectories of URA3 expression levels for a cell in the presence of low (blue) or high (red) 6AU concentrations, which alter the minimum URA3 expression level at which fitness-directed stochastic tuning can occur. We show results for a starting URA3 level [URA3]=25, with optimal fitness occurring at [URA3]=80. The initial and optimal URA3 levels are shown as gray lines. (**B**) Violin plots of the distributions of the minimum time required to reach a URA3+state (defined as [URA3]>75) in the presence of increasing concentrations of 6AU (implemented as higher thresholds of URA3 required for stochastic tuning to become active). In each case distributions reflect 50 independent trajectories simulated at each 6AU level. (**C**) Experimental validation of model predictions. Cells were grown in liquid ura-/6AU1 media (-URA) for 3–4 hr and then had the expression of fluorescent reporter proteins compared (using flow cytometry) with those of the equivalent cells grown in SC+glu (+URA) over the same time period. Values show $\log_2$ fold changes from SC+glu to ura-/6AU1; error bars show bootstrap-based 95% confidence intervals. Biological replicates performed on different days are shown side by side; the order of replicates is matched for URA3-mRuby and DHFR-GFP.

DOI: https://doi.org/10.7554/eLife.31867.010

that DHFR-GFP is also being driven by the same synthetic promoter. This again demonstrates that the tuning process occurs independently of conventional gene regulation by dedicated sensory and regulatory input.

Tuning dynamics at the single-cell level

We utilized time-lapse fluorescence microscopy to monitor the correspondence between expression of URA3-mRuby and cell division in P_{HSP12}-URA3-mRuby/P_{ADH1}-DHFR-GFP cells that initiated the tuning process. Consistent with our proposed tuning model, gene expression fluctuations that surpassed a threshold for alleviating the URA3 deficit were reinforced over long timescales and were

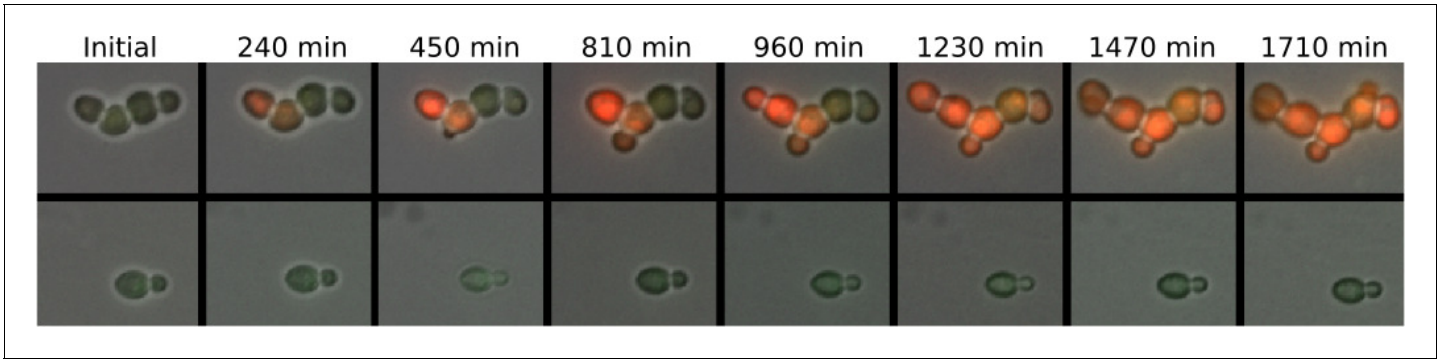


Figure 6. Sustained trans-generational inheritance of URA3-mRuby expression in tuned microcolonies. Shown are fluorescence microscopy time courses of microcolonies beginning after 12 hr of exposure to ura-/6AU5 media. A tuned colony is shown on top and a nearby untuned colony on the bottom. Fluorescence values are uniformly scaled but are not otherwise processed.

DOI: https://doi.org/10.7554/eLife.31867.011

sustained (inherited) across multiple generations as the tuned colony expanded (***Figure 6***). As expected, there is no accompanying increase in DHFR-GFP. Similar trajectories were observed for other tuning micro-colonies (***Figure 7A***). The apparently long autocorrelation time of URA3-mRuby fluctuations through the duration of a tuning trajectory is consistent with our proposed fitness feedback reinforcement mechanism. In order to quantitatively determine the timescale of gene expression fluctuations, also known as mixing times (***Sigal et al., 2006***), we utilized fluorescence-activated cell sorting (FACS) to sort a population of cells for the bottom 20%, the top 20%, and complete (mock-sorted) distribution of URA3-mRuby expression and measured the timescales over which the sorted fluorescence distributions converged to each other (***Figure 7—figure supplement 1***). For cells growing under uracil-replete conditions (SC+glu), we observed a relatively fast mixing time on the order of ~100 min (***Figure 7—figure supplement 1***; ***Supplementary file 3***). On the other hand, cells starving in ura-/6AU10 media had mixing times that ranged from 400 to 1200 min (***Figure 7—figure supplement 1***; ***Supplementary file 3***).

To determine the association of URA3-mRuby levels across generations with growth, we primed cells with 12 hr of exposure to ura-/6AU5 media and then tracked the division of tuned vs. untuned microcolonies of P_{HSP12}-URA3-mRuby/P_{ADH1}-DHFR-GFP cells over 24 hr time courses in ura-/6AU5 media. By comparing the fluorescence of cells that are about to divide with those that are not, we found that dividing cells have significantly higher levels of mRuby than non-dividing cells, whereas the separation was much smaller for GFP (***Figure 7C***). Furthermore, the URA3-mRuby levels within the tuning colony were highly heritable; as seen in ***Figure 7D***, as the indicated colony tunes and grows, cells within that colony maintain a high-mRuby state through subsequent divisions, and even their internal rankings are mostly preserved. mRuby levels in other, non-tuned microcolonies are almost uniformly lower than cells in the tuned colony. The fitness-driven optimization component of our model (***Figure 1***) further predicts that fluorescence levels should not only be heritable, but also that cells will continue to increase URA3 expression (possibly noisily) until they reach either a local optimum fitness or some biological constraint on maximum gene expression. Consistent with our expectation, we observed that the ratio of mRuby to GFP levels (the latter of which is fused to a gene whose product is not needed for growth) became steadily higher in cell lineages that had been dividing for longer (***Figure 7E***). These observations demonstrate that the level of URA3 expression is correlated with fitness, is transmitted across several generations, and shows an ongoing upward trend in tuned cells over the course of time. That last finding is particularly important because a directed increase in URA3 once a lineage begins growing is predicted by our model for fitness-directed tuning, but cannot be explained by other competing hypotheses. The images and data shown in ***Figure 7*** were taken for colonies within a single field of view of a 40x objective to ensure internal consistency in illumination and normalization, but their behavior is representative of our observations across multiple such windows. (*e.g.*, ***Figure 7—figure supplement 2***, panel A). Similar quantitative analysis from another experiment beginning directly from growth in SC+glu (instead of short-term pregrowth in ura-6AU media) is shown in ***Figure 7—figure supplement 2***, panels B-D.

Growth on –ura/6AU media does not arise from genetic mutations

It is crucial to exclude the possibility that genetic mutations underlie the observed tuning transition on –ura/6AU plates. The ongoing emergence of the tuned state in non-growing cells, over the course of many days, makes mutational mechanisms unlikely. In addition, as seen by microscopy (***Figure 3E***), no more than 1–3 cell divisions occur prior to the onset of sustained growth in a small fraction of cells.

Nevertheless, given the phenomenon of stress-induced mutagenesis in non-growing bacterial cells (***Al Mamun et al., 2012***), we wished to conclusively exclude any possibility of mutational mechanisms. To this end, we note that changes in URA3 expression occurring due to mutations should be stably heritable in the progeny of the tuned cells, which would not be expected to revert to a URA3 low state even after restoration of uracil in the media. To test the reversibility of the URA3 high state, we designed an experimental setup in which tuned colonies isolated from ura-/6AU plates were grown for varying numbers of passages in uracil-replete media (SC+glu including uracil) and then re-exposed to uracil starvation (***Figure 8—figure supplement 1***). If any genetic mutation were responsible for increasing URA3 expression in the tuned cells, the phenotype should be stable for many generations. On the other hand, stochastic tuning would predict that cells revert to a naïve state following sufficient growth in uracil-containing conditions, as they no longer benefit from URA3

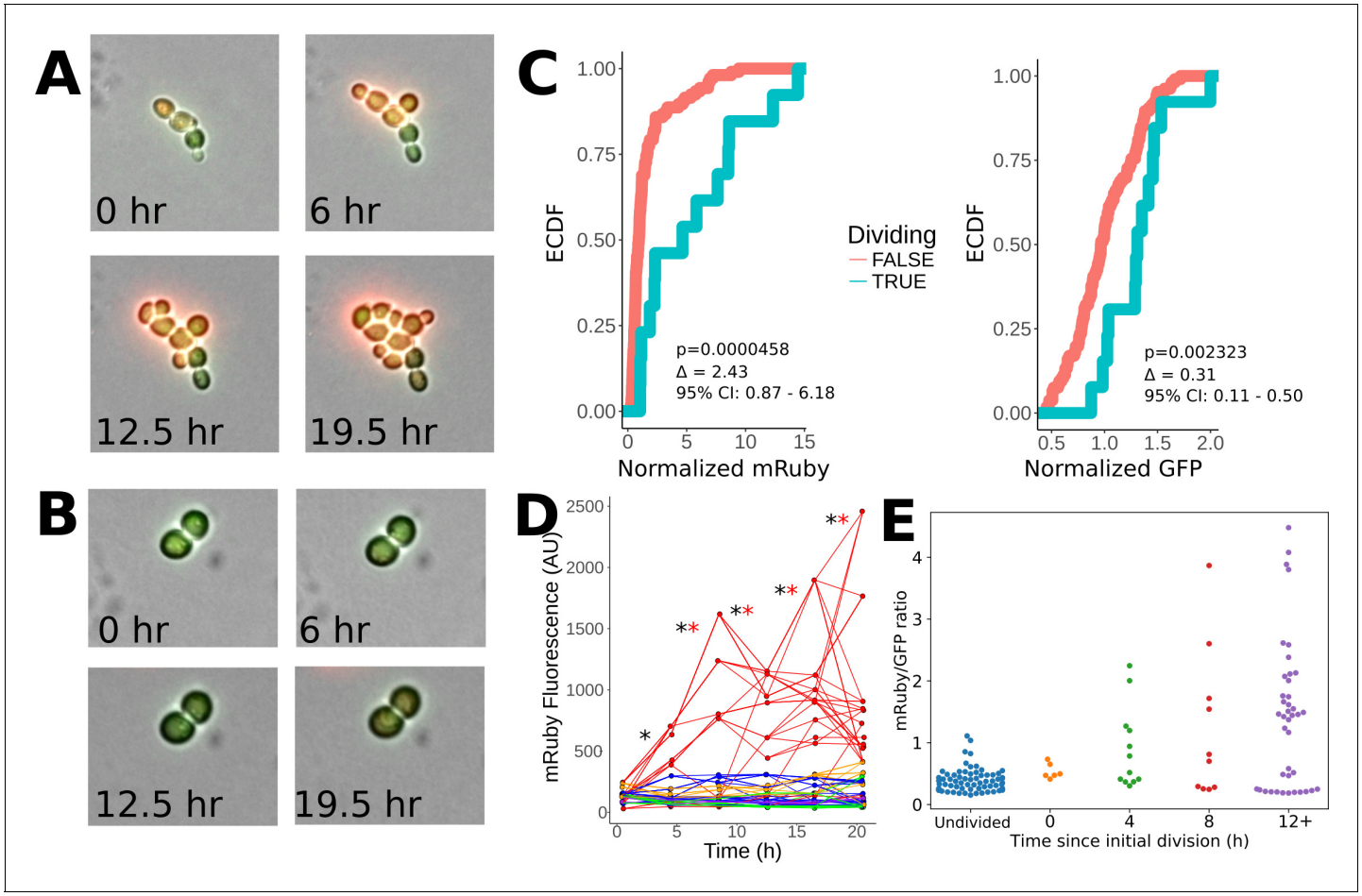


Figure 7. Heritability of elevated mRuby levels during tuning. (**A**) Formation of a microcolony over 24 hr of exposure to ura-/6AU5 media in P_{HSP12}-URA3-mRuby/P_{ADH1}-DHFR-GFP cells. GFP and mRuby are shown as transparent green and red overlays. (**B**) Snapshots equivalent to **A**) for a non-tuned colony in the same field of view. (**C**) Observed cumulative distributions (empirical cumulative distribution function; ECDF) of mRuby (left) and GFP (right) levels for cells that either do or do not divide in the timepoint following the measurement (analyzed in four-hour intervals). Values are pooled over all timepoints except the first, for five colonies growing in a single field of view. p-values arise from a Wilcoxon rank sum test applied to the shift between the non-dividing and dividing cells. Δ indicates a point estimate for the difference in fluorescence of the dividing vs. non-dividing cells, along with a 95% confidence interval (95% CI). Values shown are raw fluorescence normalized by the median value for all observations of each fluorescent protein; note the different x scales for mRuby vs. GFP. (**D**) Lineage traces showing long term propagation/inheritance of URA3-mRuby protein levels. At each specified timepoint, the average fluorescence of each cell is shown on the y axis, with lines connecting each cell to the cell(s) arising from it at the subsequent timepoint; thus, forks in the lines indicate cell division. Colors specify which of five microcolonies a given cell is a part of; only the red microcolony showed notable tuning over the course of the experiment. A black '*' is shown for each transition between adjacent timepoints for which the correlation of ranks between the timepoints in question is significant ($p<0.05$) using a Spearman correlation test, and a red '*' is shown for transitions where the same criterion holds considering only the rank ordering of cells in the red (tuned) colony (the colony shown in panel **A**). (**E**) Observed distribution of mRuby/GFP ratios depending on time elapsed since a lineage of cells began to divide. The x axis divides the cells up by the time (measured in four-hour intervals) that has elapsed since the first observed division event of an ancestor of that cell; 'Undivided' indicates cells in lineages that have not yet divided in the analyzed trajectory, and 0 hr denotes cells that will divide before the next analyzed snapshot. Note that points are plotted for each cell at each analyzed frame relative to its own growth history, and thus not all cells at a given x position necessarily arise from the same time point in the image series.

DOI: https://doi.org/10.7554/eLife.31867.012

The following figure supplements are available for figure 7:

Figure supplement 1. Mixing times of mRuby levels for growing (uracil-replete) and uracil-starved cells.

DOI: https://doi.org/10.7554/eLife.31867.013

Figure supplement 2. Heritability of elevated mRuby levels during early tuning.

DOI: https://doi.org/10.7554/eLife.31867.014

expression. As seen in ***Figure 8A***, cells with synprom-driven URA3 show reversion toward the naïve colony formation rates upon growth in (uracil containing) SC+glu media, with recovery apparent

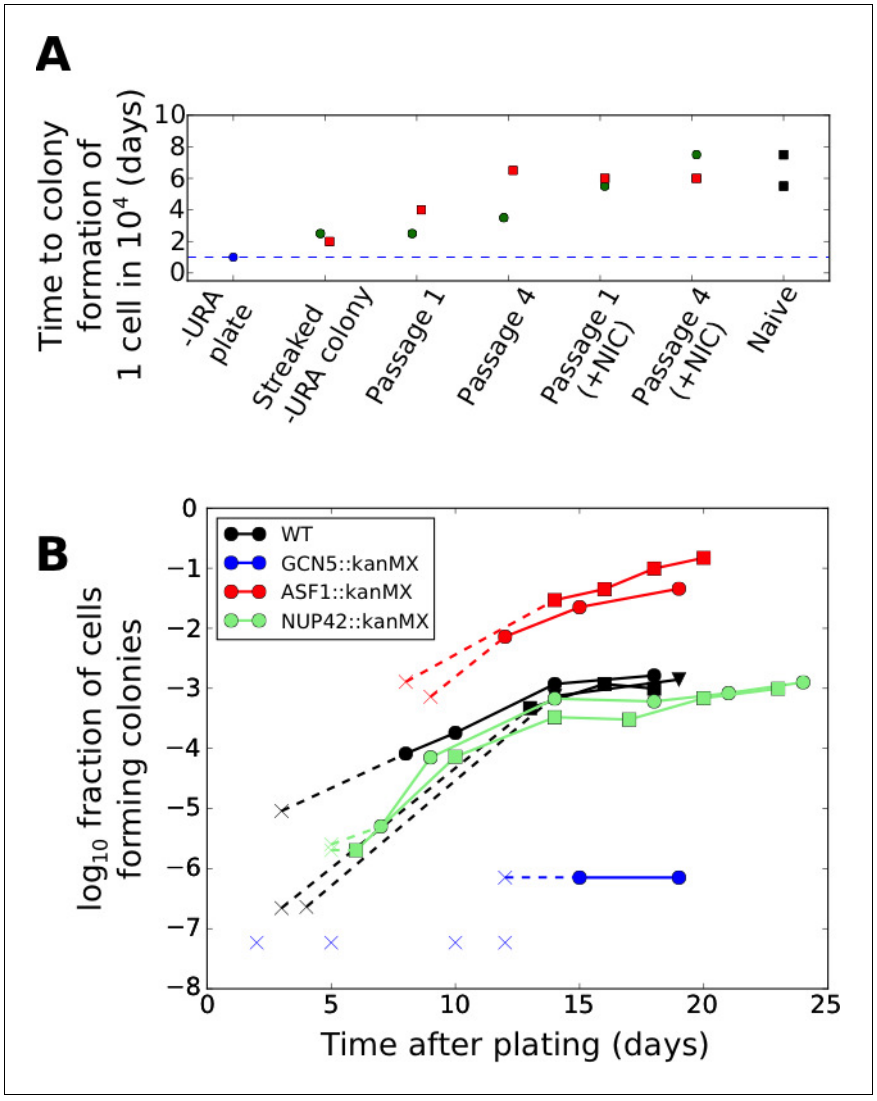


Figure 8. Effects of genetic and chemical perturbations on the efficacy of fitness-directed stochastic tuning and its epigenetic reversion. (**A**) Time courses of recovery back to the naïve state for tuned synprom-URA3-mRuby/P_{ADH1}-DHFR-GFP cells grown in either SC+glu or SC+glu with 25 mM nicotinamide added (+NIC). Extremes are shown for the colony formation times of cells never exposed to –ura conditions (Naïve) and for single colonies isolated after streaking out cells from ura-/6AU15 plates onto SC+glu (Streaked –URA colony). Colors of points indicate a single lineage beginning from a single streaked out colony picked at the first SC+glu plate stage. The cells were then repeatedly passaged in liquid SC+glu media and assessed for colony formation rates on ura-/6AU15 plates on subsequent days, as detailed in ***Figure 8—figure supplement 1***. (**B**) Colony formation rates on ura-/6AU15 plates in the presence of various genetic perturbations, assessed by colony counts from platings of selected dilutions of cells. An 'x' followed by a dashed line indicates no observed colonies and is shown at the threshold of detection from the experiment. All mutations are in a synprom-URA3-mRuby/leu2Δ0 background.

DOI: https://doi.org/10.7554/eLife.31867.015

The following figure supplements are available for figure 8:

Figure supplement 1. Recovery of cells taken from ura-/6AU plates toward a naïve state.

DOI: https://doi.org/10.7554/eLife.31867.016

Figure supplement 2. Survival of cells in –ura media in the presence of genetic perturbations.

DOI: https://doi.org/10.7554/eLife.31867.017

even after a single round of growth on an SC+glu plate, and subsequently becoming stronger with additional SC+glu passages.

To conclusively exclude mutational mechanisms, we performed untargeted whole-genome re-sequencing of a total of eight isolates with synprom-driven URA3-mRuby (four colonies from 6AU15 plates and four separate biological replicates taken after the onset of growth in 6AU5 liquid media; see Materials and Methods for details). For each case, we scanned the region within 25 kb of the LEU2 locus (where the URA3 cassettes were integrated) for mutations, since control of URA3 expression was shown in these cells to operate locally in cis (***Figure 4*** and ***Figure 4—figure supplement 2***). The results are summarized in ***Supplementary file 4***: Of the eight isolates, five show no mutations within 25 kb of the URA3-mRuby insertion, two show SNPs of unknown fitness contribution in a minority of the population, and one shows a duplication of the URA3-mRuby cassette (based on the presence of a read density that is twice the level observed elsewhere for the same chromosome). These data clearly indicate that the origin of growth-supporting URA3 expression levels in these cells cannot be reliant on a mutational mechanism, as only one of the eight cases – that with the URA3 duplication – shows a mutation at high enough levels in the population to explain the onset of growth (mutations present in less than half of the population must have arisen after one or more cells in the population had already tuned and began growing, and thus by definition could not be responsible for the initial onset of the growing state). The phenotypes caused by the sequence variants observed in populations C2 and L4 are not immediately obvious, but even if they are beneficial, their presence in a minority of cells excludes the possibility that they were responsible for the onset of tuning. Note that it should not be surprising (and, indeed, would be expected) that beneficial mutations might arise in a population once it had begun expanding in a new environment due to stochastic tuning. Our findings are consistent with a non-genetically heritable basis for the observed tuning in seven out of eight of the cases examined, as in all other growing lines, mutations near the URA3 gene were either non-existent or present only in a minority of the population.

Excluding growth-selection on the basis of pre-existing variation in URA3 expression level

A formal possibility for colony formation in a subset of the population is that growth occurs solely on the basis of pre-existing URA3 levels in cells prior to being exposed to uracil deprivation. Microscopic observations of starving cells (***Figure 3E***) argue against this possibility, as a substantial lag passes before any cells begin sustained growth. Also, colony formation continues over the course of many days (***Figure 3B–D***), demonstrating that even cells that were non-growing for a substantial time period after exposure to URA- stress can eventually grow under this condition. Nevertheless, to conclusively discount the possibility of pre-existing URA3 levels determining tuning, we sorted populations of cells on the basis of initial URA3 expression, isolated those with the highest mRuby levels (the top 0.5–1%, well outside of the main distribution of the population) and plated them. These experiments clearly showed that the ability to form colonies on ura-/6AU plates is not restricted to cells with initially high URA3-mRuby expression (***Supplementary file 5***), as the highly fluorescent cells do not form colonies on ura-/6AU plates at rates substantially higher than unsorted cells, and certainly not at a sufficiently higher rate to fully explain the observed colony formation rates. These data argue against the possibility that growth occurs only in cells that, by chance, already have high levels of URA3 expression at the time of plating (although such cells may have some slight advantage, given the nature of their initial state).

Stochastic tuning is affected by genetic perturbations to chromatin modification machinery

The proposed fitness-directed tuning mechanism relies on the capacity of local chromatin to maintain a memory of recent changes in transcription, and to modulate the transcription rate based on the fitness consequences of those changes, as conveyed by the proposed central metabolic integrator of health/fitness. We hypothesized that chromatin modification machinery may be intimately involved in these processes.

To probe the mechanistic basis of stochastic tuning, we focused on perturbations to histone acetylation/deacetylation (deletions of GCN5, SIN3, HST3, HST4), and chromatin remodeling (deletions of ASF1, ISW2, SWR1, UBP8), all of which provide potential pathways for coupling feedback from

the cell's physiological state to allele-specific modulation of chromatin and transcription (See ***Table 1*** for details). We selected these targets because of their association with genes showing particularly high levels of noise (and thus, more likely to be driven by tuning) in single-cell proteomic analysis (***Newman et al., 2006***). In our screening, homozygous replacements of HST3, HST4, SWR1, ISW2, and UBP8 with a kanMX cassette showed little effect on colony formation rates on ura-/6AU plates, and SIN3::kanMX/SIN3::kanMX strains showed severely compromised cell survival under growth-arrested conditions; all were excluded from further analysis. On the other hand, we found that genetic perturbations to the histone acetylation machinery through deletion of the key histone acetyltransferase GCN5 essentially abolished tuning, whereas deletion of the histone chaperone ASF1, in contrast, increased tuning rate by more than an order of magnitude (***Figure 8B***). At the same time, we show that the observed tuning process does not rely on transcriptional memory mechanisms grounded in chromatin localization, given the lack of effect of a NUP42 deletion (***Figure 8B***; *cf.* (***Guan et al., 2012***)).

Variations in colony formation rate are not a result of changes in viability

In interpreting our data on the effects of genetic perturbations on tuning (***Figure 8B***), it was crucial to consider the possibility that cells may lose viability at variable rates under different conditions, which could contribute to the observed differences in colony formation rates. We thus performed experiments to measure the rate of cell death in the presence of uracil starvation and compared the results with the different colony formation rates observed. As shown in ***Figure 8—figure supplement 2***, the effects of a mutation on survival and tuning rates are not significantly correlated. For example, deletion of GCN5 resulted in the nearly complete loss of stochastic tuning, deletion of NUP42 had no effect, and deletion of ASF1 substantially enhanced tuning, yet none of these mutations shows a change in survival rates during incubation in uracil-free media compared with wild type cells sufficient to explain the observed change in colony formation rate (***Figure 8—figure supplement 2***). Even for the poorest surviving strain, GCN5::kanMX/GCN5::kanMX, colony formation rates after ten days are 100−1000 times lower than wild type cells even though survival rates are lower only by a factor of ten.

Chemical perturbation of histone deacetylases inhibits the maintenance of the tuned state

Given the apparent importance of chromatin modifications in fitness-directed tuning, we also tested the effects of nicotinamide treatment (which inhibits the sirtuin class of histone deacetylases, or HDACs (***Bitterman et al., 2002***)) on reversion of the tuned cells back to a naïve state. As shown in ***Figure 8A***, we found that chemical inhibition of sirtuin HDACs by nicotinamide treatment substantially accelerated the decay of a tuned population to the naïve state, further highlighting the importance of histone modification in stochastic tuning. Combined with the data on knockout strains described above, our results suggest a central role for chromatin modifications in the establishment

Table 1. Summary of genetic perturbations tested for effects on tuning rates.

Perturbation	Direct effect	Effect on tuning
GCN5::kanMX	Deletion of histone acetyltransferase subunit (acts in ADA, SAGA, SLIK/SALSA complexes)	Inhibits
SWR1::kanMX	Deletion of H2AZ exchange factor	No effect
UBP8::kanMX	Deletion of SAGA complex de-ubiquitinase	No effect
SIN3::kanMX	Deletion of Rpd3S/L histone deacetylase components	No effect
HST3::kanMX	Deletion of Sir2-family histone deacetylase	No effect
HST4::kanMX	Deletion of Sir2-family histone deacetylase	No effect
ISW2::kanMX	Deletion of DNA translocase involved in chromatin remodeling	No effect
ASF1::kanMX	Deletion of nucleosome assembly factor	Accelerates
NUP42::kanMX	Deletion of nuclear pore complex component known to be involved in transcriptional memory	No effect

DOI: https://doi.org/10.7554/eLife.31867.018

and maintenance of the tuning process, although the molecular details cannot be discerned from these data alone.

A biologically feasible implementation of stochastic tuning

The abstract model introduced in ***Figures 1–2*** demonstrates the potential utility of fitness-directed stochastic tuning to establish adaptive gene expression states without directly sensing the external environment. In order to substantiate the biological feasibility of stochastic tuning, we implemented its critical components in a plausible simulation incorporating generic features of chromatin modification and the information flow of the Central Dogma of Molecular Biology. We therefore designed and simulated a dynamical model tracking transcription rates, transcript levels, protein levels, and histone modifications in a single cell, with parameter distributions sampled from experimental data (***Figure 9A***; see Methods for details). We incorporated the possibility of adding or removing chromatin marks that can alter the transcription rates of the associated genes. Our model incorporates two classes of marks: tuning marks (T), which link cellular fitness to transcriptional output by having mark addition rates that are a function of the recent direction of change in global fitness and current number of such marks at each promoter; and stabilizing marks (S), which are added at a rate dependent on the number of tuning marks at each promoter (***Figure 9B***). At any time, the transcriptional output of the promoter is a function of the density of both tuning marks and stabilizing marks. As such, the tuning marks provide a critical connection between changes in global fitness and transcription rates, whereas the more slowly changing stabilizing marks capture the average transcriptional output over longer timescales, enabling a more stable optimization trajectory. Both T and S chromatin marks come in two varieties: positive (activating) and negative (repressive).

Our aim was to develop a generic simulation consistent with our general knowledge of coupling between chromatin modification and transcription (***Li et al., 2007***; ***Rando and Winston, 2012***; ***Zhou and Zhou, 2011***; ***Mitra et al., 2006***). As such, the tuning and stabilizing marks described here need not correspond to any specific chemical moiety or be attributed to any particular histone modification enzyme. Modulation of enzyme activity by global fitness could be due to some as yet unknown signaling pathway or, alternatively, be dependent on known metabolic substrates or cofactors, such as acetyl-CoA and NAD+ (***Lin et al., 2000***; ***Thaminy et al., 2007***; ***Tanner et al., 1999***).

As shown in ***Figure 9C***, the detailed model is capable of stochastic tuning of a single gene which strongly impacts the fitness of the cell (as would be the case for URA3 in our experimental setup). For most randomly generated gene-level parameters, stochastic tuning results in substantially higher fitness compared to when cells undergo random fluctuations in transcription levels or when transcription is fixed at a rate appropriate for a different environment, and in most cases, tuning is able to consistently achieve near-optimal expression levels. The model is robust to variations in both the sampled biological parameters (***Figure 9C***) and the parameters of the model itself (***Figure 9D***) and can locate an optimal expression level regardless of the ratio between the initial and target protein levels (***Figure 9E***). These results demonstrate that a generic, biologically feasible implementation of fitness-directed stochastic tuning can in fact function even in the presence of the multiple layers of noise and temporal delays acting between transcription rates (at which tuning occurs) and protein levels (which dictate fitness). Note that we do not expect to find conditions where stochastic tuning is the primary mechanism of gene expression modulation for every gene in the genome, even for novel or extreme environments. Rather, we expect that the cells' hard-wired transcriptional regulatory logic exerts the primary role in the transcriptional reprogramming of the majority of genes in the genome. For its part, we expect that stochastic tuning plays the dominant role in modulating the expression of few genes/pathways that represent critical bottlenecks for fitness (for example, induction of a drug efflux pump, or repression of an enzyme that activates a pro-drug chemotherapeutic agent).

Discussion

We have described a mechanism of adaptation through fitness-directed optimization of gene expression. In numerical simulations, the proposed framework has the remarkable capacity to simultaneously tune the expression of thousands of genes, enabling optimization of fitness without directly sensing environmental parameters. The demonstration that a phenomenon consistent with fitness-directed stochastic tuning operates in *S. cerevisiae* has important implications for the

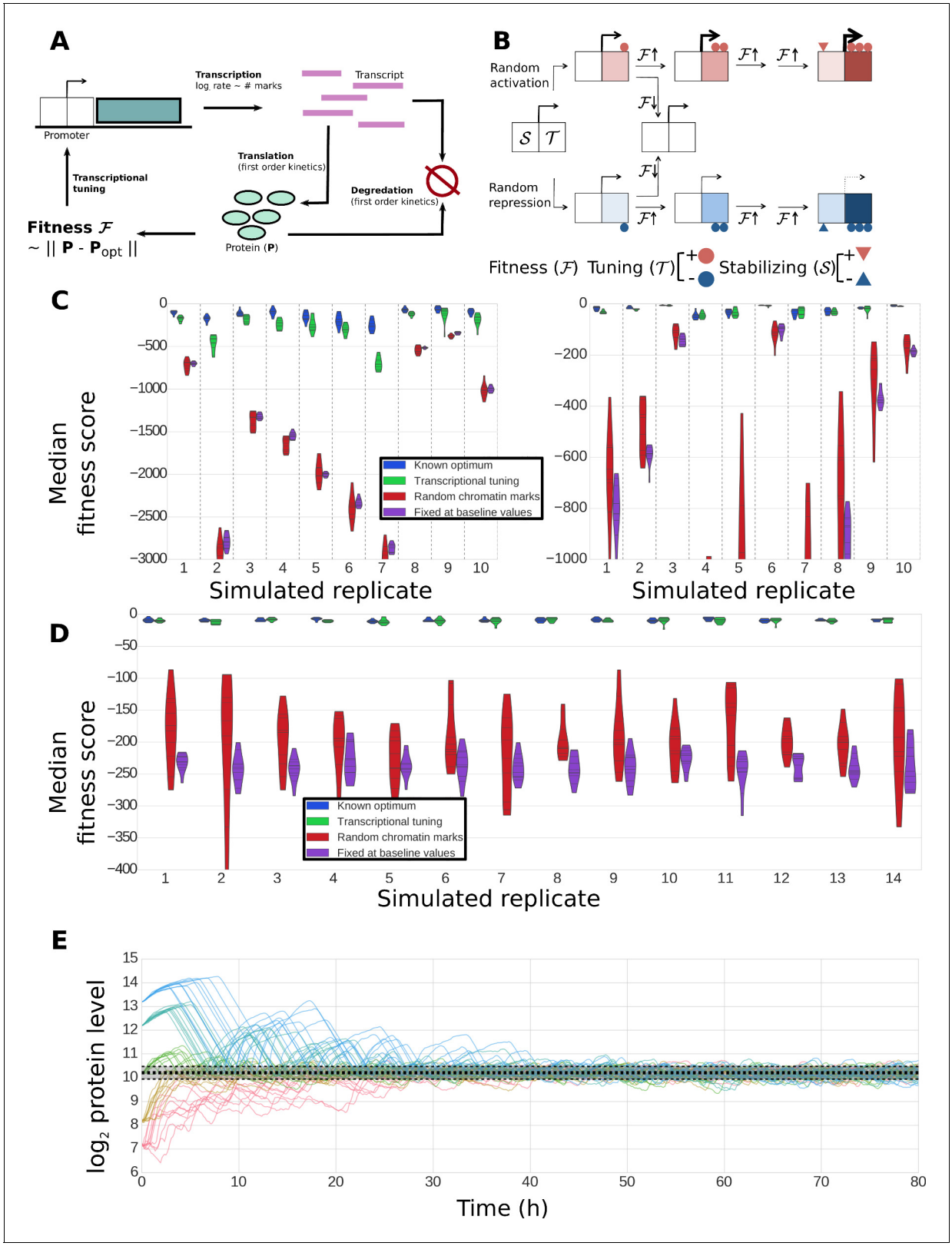


Figure 9. Construction and performance of a biologically feasible model for fitness-directed stochastic tuning. (**A**) Schematic of processes modeled in the simulation. Transcripts are produced at a rate dependent upon the state of chromatin marks at each promoter; each transcript has a fixed, gene-dependent probability of being translated at each timestep (producing a protein), and may also be degraded (again, with a gene-dependent probability). Similarly, each copy of a protein may be degraded at each timestep with a protein-dependent probability. The fitness of the system is

Figure 9 continued on next page

Figure 9 continued

calculated as the Euclidean distance between the current profile of protein counts present in the cell from a target optimum. Chromatin marks may be added or removed at each promoter at each step, as shown in panel **B**). (**B**) Logic underlying changes in tuning and stabilizing mark counts at each step. Tuning marks (T) may be added or removed at each step based on the recent history of changes in fitness, and whether each promoter currently has a net positive (activating) or negative (repressive) T count. Stabilizing marks (S) provide longer term integration by adding activating or repressive marks over time in response to the state of the tuning marks. Thus, if an unmodified promoter undergoes random addition of a positive tuning mark (top path), and that addition proves favorable, it will undergo further addition of positive T marks. If fitness continues to increase, stabilizing marks (S) will be added to stabilize its higher activity. Similar logic holds for the random addition of negative tuning marks (bottom path). In both cases, if the random T-mark perturbation proves unfavorable, the promoter will be modified in the opposite direction, in this case returning it back to its original unmodified state. (**C**) Distributions of fitness scores for a one-gene system obtained in twenty simulations using different randomly sampled biological parameters (e.g., transcript stabilities, translation rates, etc.) – these different parameter sets are the 'simulated replicates' referred to on the x axis. The median scores over the last quarter of the simulation are shown for 10 independent tuning trajectories (differing in their random number seeds). Each simulation proceeded for 300,000 steps (83.3 hr of simulated time). Different colors indicate varying methods used to control transcription rates (as shown in the legend): 'Known optimum' refers to a case where transcription rates are kept fixed at their predefined target values, 'Stochastic tuning' is the full model described in the Methods section, 'Random chromatin marks' is equivalent to the tuning model except that the direction of T chromatin mark addition is random instead of fitness directed, and 'Fixed at baseline values' shows the case where transcription rates are fixed at their initial values (intended to correspond to the environment that the cells were in prior to the onset of stress exposure). Dashed vertical lines group simulations performed with identical parameters. On the left axis we show ten sets of simulations where the target transcription rate was eight-fold higher than the starting rate, and on the right axis simulations where the target transcription rate was eight-fold lower than the starting rate. (**D**) Robustness of tuning against changing model parameters. Violin plots are defined as in panel **C**), but in this case show the distributions of fitness scores observed under variations of the model parameters (*e.g.*, magnitude of individual tuning and stabilizing marks) for a single, randomly chosen set of gene-specific parameters. Plotted are the median fitness scores over the last quarter of each simulation, using either our central 'baseline' parameters for all model parameters (leftmost replicate; see ***Supplementary file 8***), or twofold changes (up or down) of each editable parameter in our model. (**E**) Tuning performance of a single gene matching a wide range of biological challenges. For a fixed set of biological parameters (see Materials and Methods), we performed 10 simulations each where the initial transcription rates were off from the target rate by a factor of 2^3, 2^2, 0, 2^{-2}, and 2^{-3}, running in order from blue to red. A strong dashed black line shows the median obtained from the last quarter of a long (3 million step) simulation with transcription rates fixed at their optimal values; the shaded region shows the extent of a region encompassing 95% of the timepoints observed in that window. Regardless of initial conditions, the protein level approaches the optimal value and then stably oscillates around it, with amplitudes similar to those observed in the control simulation with target transcription rates.

DOI: https://doi.org/10.7554/eLife.31867.019

adaptation of eukaryotic microbes to novel or extreme environments where their genetically encoded regulatory networks become inadequate. However, we speculate that stochastic tuning operates in parallel with conventional regulation even in frequently encountered environments. Indeed, hard-coded sensory and regulatory networks are unlikely to have the encoding capacity to optimally respond to every conceivable subtle change in the environment—even within the native habitat. We therefore favor a model in which dedicated regulatory networks quickly move the system to a state reasonably well matched to a given condition, and stochastic tuning subsequently optimizes expression to achieve a more precisely adapted state for every individual encounter.

The ability to discover optimal gene expression states through a stochastic fitness-directed search may have provided significant advantage to early eukaryotic microbes. Microorganisms have evolved stochastic search strategies in other contexts. Indeed, the proposed stochastic tuning mechanism is reminiscent of the biased random walk phenomenon in bacterial chemotaxis, where stochastic transitions in the rotation of the flagellar motor are biased towards the direction that increases chemoattractant signaling over time (***Macnab and Koshland, 1972***). Detailed molecular mechanisms of chemotaxis have been revealed over the course of the last few decades, demonstrating the versatility of molecular processes in implementing rather complex computations (reviewed in (***Sourjik and Wingreen, 2012***)). Although our main focus here has been on establishing the phenomenology of fitness-directed stochastic tuning, we have already identified some critical components. In particular, histone acetylation/deacetylation (via GCN5 and sirtuins) seem to play a critical role, as deletion of GCN5 almost entirely abolished tuning. This is consistent with the high degree of intrinsic noise exhibited by the genes that are regulated by the SAGA complex, in which GCN5 is the catalytic subunit (***Newman et al., 2006***). Previous work has shown that increased transcriptional noise is beneficial for adaptation to acute environmental stress (***Blake et al., 2006***). Interestingly, however, early work demonstrated that deletion of GCN5 further increases expression noise in the context of the PHO5 promoter (***Raser and O'Shea, 2004***).

Taken together, these data suggest that stochastic tuning is not driven by noise alone; rather we support a model in which the proper integration of noise, transcriptional memory, chromatin modification, and cellular-health feedback work together to implement a directed search mechanism to drive the expression level of individual genes to levels that maximize the overall health of the cell. Indeed, histone modification is tightly coupled with gene expression. Co-transcriptional histone modification can store recent memory of transcriptional activity (**Li et al., 2007**; **Rando and Winston, 2012**) and histone modification can, in turn, affect transcription rate (**Stasevich et al., 2014**). There has been a longstanding debate on the functional significance of this reciprocal coupling. Our model and results help to unify these phenomena and support their functional relevance as requisite components of a stochastic tuning-based cellular adaptation framework.

We note that our experimental setup for demonstrating stochastic tuning has superficial similarities to a series of experiments performed in *S. cerevisiae* by the Braun lab, in which they sought to determine whether glucose-driven repression of the GAL1 promoter could be overcome to allow expression of a HIS3 construct in glucose-containing media (**Stern et al., 2007**; **Stolovicki et al., 2006**). While the authors observed consistent emergence of growth in a large fraction of cells that they initially noted could be attributed to either genetic or epigenetic mechanisms (**Stolovicki et al., 2006**), subsequent analysis has shown that in that experimental system, genetic mutations are the primary mechanism of adaptation, possibly driven by hypermutability of the genes involved in the response of interest (**David et al., 2010**; **Moore et al., 2014**; **David et al., 2013**). These mutational mechanisms stand in clear contrast to the rapidly reverting epigenetic stochastic tuning observed in our experiments.

In addition to perception of environmental parameters, cells also possess a variety of hard-wired homeostatic mechanisms sensing and responding to internal parameters, optimizing resource allocation in response to parameters such as growth rate (**Klumpp et al., 2009**; **Klumpp and Hwa, 2014**; **Brauer et al., 2008**; **Barenholz et al., 2016**; **Keren et al., 2013**) and metabolite/nutrient pools (**Potrykus et al., 2011**; **Broach, 2012**). However, while these mechanisms allow cells to sense their internal state, they still reflect specific evolved responses to alter resource allocation and gene expression in a *predefined* way in response to stress, standing in contrast with the ability of stochastic tuning to conduct a search and discover arbitrary gene expression states that are adaptive under extreme and unfamiliar environments.

The widely varying tuning rates for different promoters (**Figure 3B–C** and **Figure 3—figure supplement 1**) clearly indicate that sequence features can influence tuning efficacy. By design, all but one promoter driving URA3 in our experiments contained a TATA box, which has been linked to high intrinsic noise (**Newman et al., 2006**), condition-specific expression variability (**Tirosh et al., 2006**) and reliance on chromatin-mediated regulation (**Tirosh et al., 2008**; **Basehoar et al., 2004**). Indeed, replacement of the (TATA-containing) P_{SAM3} derived sequence in synprom with a similarly generated sequence from the TATA-free P_{ARF1} promoter substantially reduced tuning rates under the conditions tested (**Figure 3—figure supplement 1**). We also note that when we performed experiments similar to those described above with the repressed natural promoter P_{GAL1}, we observed dramatically lower rates of colony formation (less than 1 in 10^7), and those colonies that did form appeared to be non-reverting genetic mutants (data not shown). Exploring the full importance of transcriptional noise for tuning efficiency, as well as that of other features such as propensity for nucleosome positioning, will be important in future work.

Fitness-directed stochastic tuning requires feedback of the global state of health to every promoter in the genome. The dependence of many histone modification enzymes on metabolic intermediates and cofactors (e.g., NAD+ for the sirtuin family of histone deacetylases (**Lin et al., 2000**; **Thaminy et al., 2007**); SAM for histone methyltransferases (**Luka et al., 2009**), and acetyl-CoA for histone acetyltransferases (**Tanner et al., 1999**)) provides support for potential direct feedback of global fitness-related parameters to the epigenome (**Katada et al., 2012**; **Kurdistani, 2014**), and indeed we showed that chemical manipulation of sirtuin activity had substantial effects on retention of epigenetic memory. These enzymes may very well serve as distinct channels of health-related information utilized by stochastic tuning. In this regard, chromatin itself may function as a global health integrator, with histone modifications and their effect on gene expression being highly contingent on the current trajectory of cellular fitness. Alternatively, cells may utilize a single global health integrator (such as the mTOR system) as hypothesized in our idealized model. The mTOR pathway integrates diverse parameters of internal health including energy, nutrient availability, and cellular

stresses (**González and Hall, 2017**). Intriguingly, the mTOR pathway has recently been shown to regulate histone acetylation states through a variety of mechanisms (**Chen et al., 2012**; **Workman et al., 2016**)

Fitness-directed stochastic tuning has important implications for gene regulation. Beyond a potentially widespread mechanism of cellular adaptation, stochastic tuning brings together seemingly unrelated phenomena under a unifying conceptual framework. These are areas of study at the frontier of genetics and biochemistry, including stochastic gene expression, transcriptional memory, and metabolic modulation of epigenetic states. Stochastic tuning may have initially evolved as a mechanism for adaptation of single-cell eukaryotes to extreme environments. However, once available, it may have found additional utility as a versatile mechanism for controlling and fine-tuning gene expression in the context of physiological and developmental processes in metazoans. This is consistent with the evolutionary arc of an ancient set of molecular mechanisms that now serve as key mediators of differentiation (**Álvarez-Errico et al., 2015**; **Ziller et al., 2015**; **Meissner, 2010**). Exploring this possibility represents an important area for future research. Optimization of cellular health through the fitness-directed stochastic tuning mechanism may also play an important role in allowing cancer cells to survive and thrive in a variety of microenvironments unfamiliar to their evolved regulatory networks, and in the face of extreme challenges imposed by chemotherapy and radiation. Indeed, stochastic tuning may underlie the epigenetically mediated metastatic potential and chemotherapy resistance observed in a variety of cancer types (**Wu and Roberts, 2013**; **Perez-Plasencia and Duenas-Gonzalez, 2006**; **Lv et al., 2016**; **Li et al., 2015**; **Borley and Brown, 2015**; **Bonito et al., 2016**; **Shaffer et al., 2017**). Our observations support the existence of a fitness-directed tuning process that operates at the level of transcription. However, in principle, tuning could also occur at any point along the hierarchy of gene expression where noise, memory, and feedback of global fitness can drive the activity of gene products towards levels that optimize the overall health of the cell. In particular, searching for evidence of tuning at the level of translation would be an important focus for future research.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
gene (*Saccharomyces cerevisiae*)	URA3	NA	YEL021W	
gene (*Entacmaea quadricolor*)	mRuby	DOI: 10.1371/journal.pone.0004391		
gene (*Aequorea victoria*)	GFP	DOI: 10.1038/nbt1172		Codon optimized for *S. cereivisiae*; sequence available as **Supplementary file 3**
genetic reagent (*S. cerevisiae*)	P_{HSP12}	NA		Promoter region upstream of YFL014W
genetic reagent (*S. cerevisiae*)	P_{ADH1}	NA		Promoter region upstream of YOL086C
genetic reagent (*S. cerevisiae*)	P_{RGI1}	NA		Promoter region upstream of YER067W
genetic reagent (S. cerevisiae)	synprom	This paper		Synthetic promoter sequence. See Supplementary Material for complete sequence, and methods for details of construction
genetic reagent (S. cerevisiae)	GCN5::kanMX	PMID: 10436161		Knockout cassette obtained from the yeast knockout collection strain
genetic reagent (S. cerevisiae)	ASF1::kanMX	PMID: 10436161		Knockout cassette obtained from the yeast knockout collection strain
genetic reagent (S. cerevisiae)	NUP42::kanMX	PMID: 10436161		Knockout cassette obtained from the yeast knockout collection strain
strain background (*S. cerevisiae*)	BY4743	PMID: 9483801		
chemical compound, drug	6-azauracil	ACROS Organics	Product code 153970050	Stock solution 10 mg/mL in 1 M ammonium hydroxide

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
chemical compound, drug	Nicotinamide	Sigma	Product number N0636	Stock solution 1 M in water; filter sterilized
software, algorithm	tuning_simple	This paper		Octave implementation provided as ***Source Code 2***
software, algorithm	tuning	This paper		Python implementation provided as ***Source Code 3***

Media and strains

For routine growth of strains, we used YPD broth (10 g/L yeast extract, 20 g/L peptone, 20 g/L dextrose) or YPD agar plates (YPD broth +20 g/L Bacto agar). We used standard recipes based on SC +glucose (SC+glu) (***Kaiser et al., 1994***) for all physiological experiments. SC/loflo refers to SC made with low fluorescence yeast nitrogen base (US Biologicals). In the case of SC+glu, we used dropout supplement powders interchangeably from ForMedium (DSCK012) and US Biologicals (D9515), although they differ slightly in the concentrations of adenine and *para*-amino benzoic acid supplied. SC+glu derivatives lacking particular nutrients are specified as SC+glu-NUTRIENT; e.g., SC+glu-ura for SC+glu lacking uracil. We also refer to the commonly used mixture of SC+glu-ura with 6-azauracil added as ura-/6AU*i*, where *i* is the final concentration of 6AU in microgram/mL. The agar for all plates used in physiological experiments was either Noble agar (Difco) or quadruple-washed Bacto agar. For the removal of the GAL-GIN11 cassette in counter-selections (see below), cells were plated on YPGA agar plates (10 g/L yeast extract, 20 g/L peptone, 20 g/L galactose, 20 g/L agar, 100 microgram/mL ampicillin). All growth was at 30°C; liquid phase growth included shaking at 200–220 rpm in an Innova 42 incubator (New Brunswick).

As diagrammed in ***Figure 3A***, we constructed two classes of insertion cassettes. Each follows the pattern of having a promoter, a functional reporter protein fused to a fluorescent protein, and then ends with a CYC1 terminator. For URA3, the native sequence from *S. cerevisiae* was used, with the exception of one silent SNP and an A160S mutation that does not appear to alter enzyme function. The red fluorescent protein mRuby is described in (***Kredel et al., 2009***). For DHFR, we used murine DHFR from pSV2-dhfr (***Subramani et al., 1981***) with an L22R mutation making it methotrexate-resistant (***Simonsen and Levinson, 1983***). GFP refers in all cases to superfolder GFP (***Pédelacq et al., 2006***) codon-optimized for *S. cerevisiae* using web-based tools from IDT (Integrated DNA Technologies); see ***Supplementary file 3*** for the corresponding nucleotide sequence. In each case, the reporter and fluorescent protein were separated by a short A/G/S containing linker. All constructs were cloned in bacterial hosts using pBAD-derived plasmids; separate plasmids were constructed with each promoter of interest downstream of a region homologous to the upstream target site in the *S. cerevisiae* genome, and URA3-mRuby-cyc or DHFR-GFP-cyc upstream of a region homologous to the downstream target site in the *S. cerevisiae* genome. All constructs were chromosomally integrated at the leu2Δ0 locus of our yeast strains. Double-stranded DNA for transformation in yeast was then generated by first amplifying the promoter and reporter constructs separately, using primers yielding 20–40 bp overlaps; we then used crossover PCR to generate the complete construct of interest and subsequent amplification to generate a sufficient quantity for transformation. All PCR used for strain construction was performed using Q5 high fidelity polymerase (NEB); routine PCRs for strain validation were instead performed using OneTaq or Taq polymerase (NEB).

Promoters for ADH1, HSP12, and RGI1 were cloned from our wild type strain (BY4743 or its haploid progenitors BY4741/BY4742) and included the entire region from 1700 to 1800 bp upstream of the start codon to the base immediately prior to the start codon. The ADH1 promoter was selected as a classic constitutive promoter (***DeMarini et al., 2001***); HSP12 and RGI1 were chosen as they show high variance in expression between conditions (***Tirosh et al., 2009***; ***Tirosh et al., 2006***), a characteristic expected to be favorable for stochastic tuning. Synprom was designed in two stages: the bulk of the DNA is a 600 bp random sequence generated using a Markov model to match the trinucleotide frequencies present across all natural *S. cerevisiae* promoters. To this sequence we appended the 200 bp immediately prior to the start codon of SAM3, to provide native transcription and translation start sites. The resulting sequence was then modified to remove all recognizable

binding sites for yeast transcription factors (TFs) as follows: we used the set of position weight matrices and match thresholds in ScerTF (**Spivak and Stormo, 2012**) to identify all recognizable TF binding sites in the promoter, and randomized the sequences of only those regions and their immediate surroundings until no recognizable TF binding sites remained. The resulting perturbed sequence is given as **Supplementary file 1**. The required sequences were synthesized as gBlocks from Integrated DNA Technologies and combined via Gibson assembly (**Lartigue et al., 2009**).

All yeast strains were derived from BY4741 or BY4742 (**Brachmann et al., 1998**), which includes a complete deletion of the URA3 ORF (BY4741: Mat **a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0; BY4742: Mat α his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0). Insertions of URA3 or DHFR fusion proteins were always at the leu2Δ0 locus unless otherwise noted. To facilitate consistent insertion, we replaced the leu2Δ0 allele of BY4741/BY4742 with a LEU2-GAL-GIN11 cassette (**Akada et al., 2002**), which allows growth in leucine-free media but inhibits growth in the presence of galactose. We note that at least in our copy of the BY474x strains, the leu2Δ0 deletion runs only from ChrIII:84799—ChrIII:93305, rather than extending to position 93576 as annotated. Nevertheless, the deletion is sufficient to remove the entire leu2 open reading frame.

Strains containing the fusion proteins were constructed by transforming the LEU2-GAL-GIN11 containing cells with appropriate double-stranded oligos (see above) and selection on YPGA plates, allowing replacement of the LEU2-GAL-GIN11 cassette with the desired insert. Insertions were confirmed by PCR product sizing. Diploid strains were derived by mating one BY4741-derived (mat **a**) strain with one BY4742-derived (mat α), and subsequently plating on SC+glu-lys-met or SC+glu-lys-met-cys. All transformations were carried out using the LiAc-PEG-ssDNA method (**Gietz and Woods, 2002**).

Knockout strains were generated by beginning from appropriate haploids containing either a leu2::promoter-URA3 or leu2::promoter-DHFR construct or simply leu2Δ0, amplifying an appropriate kanMX knockout cassette from the corresponding strain in the *S. cerevisiae* gene deletion collection (**Giaever et al., 2002**), and selecting on YPD+G418 plates. We confirmed the presence of kanMX at the appropriate site and absence of the native gene by PCR. Diploid knockout strains containing appropriate deletions and a URA3-mRuby insertion at leu2Δ0 were generated by mating these haploids as noted above.

Colony formation assays

Experiments showing colony formation rates over time all follow a common formula. Cells were grown overnight in SC+glu media, and then in the morning back-diluted 1:200 into fresh, pre-warmed SC+glu. The cells were grown for four to five hours at 30°C with shaking and then pelleted, washed once with 25 mL deionized (DI) water, pelleted, washed with 1 mL water, pelleted, and resuspended in 1 mL water. Specified dilutions were made in DI water from this final cell suspension.

Cells were then either plated on full plates at pre-chosen dilutions (100 microliters of an appropriate cell suspension), or a dilution series was spotted onto appropriate agar plates (10 microliters per spot). Plates were imaged and counted every 1–2 days for the duration of the experiment (lasting between a few days and weeks, depending on the experiment in question). Plates were wrapped in parafilm after ~3 days to minimize drying. Plating was performed identically on SC+glu plates (to establish the number of cells being plated) and plates containing one or more test conditions (e.g., ura-/6AU).

Cells were counted either directly from the plates or from stored digital images. Direct plate counts were done manually for all visible colonies; for those counted from saved images, we imposed a minimum size threshold of 0.2 mm in diameter (rounding up to the nearest pixel). Times for counts were rounded to the number of days since plating.

Death rate assays

To determine the survival rates of cells undergoing uracil starvation in the presence of various other perturbations, we measured the death rates of cells lacking any copy of URA3 in SC-ura+glu media. Cells were pregrown and washed as described above for plating assays, but then resuspended in liquid SC-ura+glu media and incubated at 30°C. Aliquots were regularly removed and spotted on SC +glu plates to determine the number of viable colonies. Survival rates are for leu2Δ0 homozygotes

(the original BY4743 diploid, possibly with a homozygous deletion of a specified gene) with no available copy of URA3.

Flow cytometry

Cells were analyzed by flow cytometry on an LSR Fortessa (Becton Dickinson) at the Columbia University Microbiology and Immunology Flow Cytometry Core Facility or University of Michigan Flow Cytometry Core. Cells to be used in these experiments were initially prepared and washed following the same pregrowth procedure as given above for colony formation assays, except that growth was in low fluorescence SC/loflo media instead of SC. A flask containing 25 mL of prewarmed media (generally ura-/6AU5 made from an SC-ura/loflo base) was then inoculated with 200 microliters of the cell suspension, and cells were grown with shaking at 30˚C. Subsequent data acquisition varied depending on the experiment to be performed.

For the long time courses shown in ***Figure 4*** and its supplement, for an initial timepoint, 200 microliters of the washed cell suspension were combined with 500 microliters of 2x PBS/E (1x PBS with 10 mM EDTA added), 290 microliters DI water, and 10 microliters of flow cytometry counting beads (Invitrogen CountBright beads). At subsequent timepoints, snapshots were taken by combining 490 microliters of the growing cells, 10 microliters counting beads, and 500 microliters 2x PBS/E. In either case, cells were run on the Fortessa, with signals recorded for forward and side scatter, mRuby (using the Texas Red laser/filter set), and GFP (using the FITC laser/filter set).

Data were analyzed using the flowCore and flowViz modules of R (***Ellis et al., 2006***; ***Ellis et al., 2009***). Beads and cells were first identified based on their forward scatter and side scatter (FSC/SSC) values (using permissive gates that capture the vast majority of each population) and fluorescence (beads were required to show very high fluorescence). For each growth phase (exponential in SC+glu, starving in ura-/6AU, growing in ura-/6AU), we obtained empirical autofluorescence corrections by analyzing populations in a similar growth state lacking the fluorescent tag on URA3. Guided by exploratory analysis, we fit a linear model for starving cells predicting mRuby and GFP autofluorescence as a function of the observed forward and side scatter, and used constant autofluorescence values characteristic of each of the two growing phases (obtained from cells with no fluorescent protein in a similar physiological state, either uracil-starved or undergoing stochastic tuning-driven growth). During analysis of liquid phase fluorescent populations (shown in ***Figure 4*** and its supplement), the predicted autofluorescence values were subtracted from the observed value; in these cases, an additional gate was applied to remove events with very low forward scatter values, which had a very high variance in fluorescence and were well below the size of the main population.

For the use of FACS followed by plating to test the colony formation rates of highly fluorescent cells, cells were prepared as described above, sorted using a BD FACSAria, and then subsequently plated in equal quantities on SC+glu and ura-/6AU15 plates.

For the short timescale tuning data shown in ***Figure 5C***, the cells were grown for 3–4 hr side by side in SC/loflo + glu and –ura/loflo/6AU1 media, and then placed on ice and run directly on the flow cytometer. For each biological replicate (performed on different days), we grew leu2::synprom-URA3-mRuby/leu2::synprom-DHFR-GFP and nonfluorescent leu2::URA3/leu2Δ0 cells in parallel to allow direct comparison of the observed fluorescence levels. Analysis was performed separately for each biological replicate. We first normalized all fluorescence signals by the FSC-A signal raised to the power of 1.5, which we found empirically to be an effective correction removing most of the dependence of the fluorescence on cell size. Next, a mapping of FSC signals to expected autofluorescence on each channel was fitted using the R loess function (with default parameters), and the expected autofluorescence subtracted from the observed value for each cell to yield what we refer to as the blanked fluorescence. We then calculated and compared the changes in the median blanked fluorescence of the populations for the same cells grown in SC+glu vs. ura-/6AU1 media. Confidence intervals were calculated by bootstrapping with 200 bootstrap replicates.

Whole genome sequencing

Cells for whole genome sequencing were taken directly from the growth condition of interest (ura-/6AU15 plate or ura-/6AU5 liquid media) and flash frozen in 15% glycerol or 1x TES (10 mM Tris, pH 7.5; 10 mM EDTA, 0.5% SDS). One reference sample grown under unselective conditions was taken for each starting strain to use as a baseline. Genomic DNA was isolated using a YeaStar Genomic

DNA kit (Zymo Research) according to the manufacturer's instructions. Samples were then barcoded and prepared for sequencing using a Nextera XT kit (Illumina, Inc.) and sequenced as part of a pooled library on a NextSeq (Illumina, Inc.).

Sequencing reads were clipped to remove adapters and commonly observed artifactual end sequences with cutadapt (**Martin, 2014**), and then further trimmed using Trimmomatic 0.30 (**Bolger et al., 2014**) to remove very low quality (<3) end bases, retain only the portion of the read with a quality score above 15 in a four base sliding average window, and remove reads less than 10 bp long. Surviving trimmed reads were then aligned to the reference genome using Bowtie 2.1 (**Langmead et al., 2009**); the reference genome was constructed from the *S. cerevisiae* S288c genome (GenBank BK006934 – BK006949), deleting the URA3 ORF and inserting the sequence for the appropriate URA3 and DHFR constructs in separate copies of chromosome III at the LEU2 locus. Read data used in this analysis are available from the Short Read Archive under accession SRP117724.

After alignment, mutational calls and read depths were obtained using the mpileup and depth modules of samtools 0.1.18 (**Li et al., 2009**), respectively. Reads for called variants within 25 kb of the insertion site were examined manually and compared to the sequenced parental strain; validated variants are listed in **Supplementary file 4**.

RNA isolation

RNA was isolated using an adaptation of the hot acid phenol method (**Collart and Oliviero, 2001**). Cells for RNA isolation were grown under appropriate conditions (either in liquid phase or on agar plates), and then snap-frozen in 1x TES (10 mM Tris, pH 7.5; 10 mM EDTA; 0.5% SDS) and stored below −70˚C. Snapshots of 200 to 600 microliters were taken from growing liquid phase cultures, whereas from agar plates we harvested 1–20 colonies of <0.5 mm diameter taken from the same plate as each biological replicate. RNA was isolated by rapidly thawing the cell suspension and mixing 1:1 with a 5:1 acid phenol:chloroform solution, then incubating 60 min at 65˚C with occasional vigorous vortexing. The solution was then chilled on ice for 5 min, and centrifuged 5 min at 16,000 x g at 4˚C. The aqueous phase was mixed 1:1 with additional acid phenol:chloroform, chilled, and centrifuged as before. The aqueous phase was then mixed 1:1 with a 24:1 chloroform:isoamyl alcohol solution, and centrifuged 5 min at 4˚C. The resulting aqueous phase was transferred to a fresh tube and combined with 1/10 vol 3 M sodium acetate, 2 volumes of 1:1 ethanol:isopropanol, and 1/800—1/200 vol Glycoblue (Ambion), and then precipitated for at least 1 hr at −20˚C and then at least 1 hr at −80˚C. RNA was recovered by centrifuging 15 min at 16,000 x g at 4˚C, washed with ice cold 75% ethanol, spun an additional 5 min, and then air-dried and resuspended in RNAse-free water. The samples were then further purified using a Zymo RNA clean and concentrator five according to the manufacturer's instructions, including an on-column DNase digestion.

Quantitative RT-PCR

Total RNA was purified from cells in the desired growth condition using the hot acid-phenol procedure described above. cDNA pools were generated for each sample using random hexamer-primed reverse transcription with Protoscript II (New England Biolabs) following the manufacturer's instructions. cDNA pools were used directly in qPCR reactions without further purifications, assembling reactions using iTaq Universal SYBR Green Supermix (BioRad) following the manufacturer's instructions, in GeneMate PCR plates. Plates were sealed with Microseal 'B' adhesive film (BioRad) and run on a BioRad CFX96 detection system. C_t values calculated by the instrument software were then exported for subsequent analysis. All isolated RNA was quantified on a Bioanalyzer (Agilent) and found to have an RIN >= 6.8.

For comparison of URA3 and DHFR expression, we calculated separate ΔC_t values for each qPCR run replicate by taking the median of all technical replicates from that run. Values plotted in **Figure 4—figure supplement 2** reflect ΔC_t data from 1 to 2 technical replicate wells on each of two to four separate, independently assembled runs; we plot the median of day-wise data points for each separate biological sample. Primer locations and sequences are given in **Supplementary file 7**. We performed a no-reverse transcriptase control reaction for each sample to ensure that DNA contamination did not contribute to the observed signal (data not shown).

qRT-PCR data were analyzed using a Bayesian hierarchical model treating the ΔC_t value between the URA3 and DHFR primers as follows:

$\Delta C_t(\text{sample,day}) \sim T(\mu_s(\text{sample}), \sigma_{rep}, \nu_{rep})$

$\mu_s(\text{sample}) \sim T(\mu_c(\text{class}), \sigma_c(\text{class}), \nu_{bio})$

Parameters not otherwise specified were assigned appropriate uninformative priors. Here 'sample' refers to a single biological sample and 'class' to a single growth condition. The key parameter of interest is μ_c for each class of cells under study, the overall average URA3:DHFR difference for cells grown under that condition. We fitted the model using JAGS (***Plummer, 2003***), and then report credible intervals and other inferences from the posterior distribution on μ_c. Each of the ΔC_t(sample, day) values used the median across 1–2 technical replicates for each primer pair.

Cell count data analysis

Data were analyzed using custom-written python and R scripts. Source code for the nontrivial analysis of flow cytometry data is provided as ***Source code 1***. Uncertainties for cell counts (shown in plating and flow cytometry data) were calculated by treating each observed count as a Poisson random variable; using Bayesian inference with the Jeffreys prior (***Jeffreys, 1961***), the posterior distribution for the rate parameter I (the concentration of cells) is given in closed form by

$I \sim \text{Gamma}(0.5 + \sum_{i=0}^{n} i_n, n)$

Where n is the number of observations and the i_n are the observed counts. Error bars then indicate a central 95% credible interval for I given the observed data.

Recovery experiments

Experiments to examine the reversion of tuned colonies toward a naïve state were performed as shown in ***Figure 8—figure supplement 1***. Single colonies from a ura-/6AU15 plate were streaked out onto SC +glu and allowed to grow. From that plate, single colonies were again picked and underwent repeated passages in liquid media; each 'passage' refers to a 200-fold dilution, which is then allowed to grow for 48 hr (96 hr for the very first transfer). Cells were also taken for plating from the original ura-/6AU15 plate, the first SC +glu plate stage, and several subsequent time points during liquid culture. Cells taken from plates were immediately diluted in water and spotted on SC +glu and ura-/6AU15 to track colony formation rates; cells taken from liquid passages were streaked out on SC +glu plates prior to use in spottings, in order to obtain a consistent physiological state. Plots for 'naïve' cells refer to cells treated identically, except that they had initially been grown on SC +glu plates instead of ura-/6AU15 plates. Recovery was assessed based on the amount of time required for 1 in 10,000 cells spotted on the new ura-/6AU15 plate to form countable colonies (using linear interpolation of colony counts between observed data points); in the event that one dilution yielded no colonies passing our size threshold, but the next (10-fold more concentrated) spot gave an uncountable haze, we assigned a count of 1 to the more concentrated spot.

Numerical simulations

The numerical simulations shown in ***Figures 2*** and ***5*** were performed by implementing the model described in the text using the Matlab programming language and simulated using Matlab (Mathworks, Inc.) or GNU Octave version 3.8.1 (***Eaton et al., 2009***), with qualitatively equivalent results obtained in either case. All simulations were performed using the same initial conditions (but different random seeds, for the sampling shown in ***Figure 5***). Octave code implementing this model is provided as ***Source code 2***.

The physiological tuning model employed for ***Figure 9*** and the accompanying text was implemented in python, and simulated using python 2.7.6, making heavy use of the numpy (***Svd et al., 2011***) and scipy (***Jones et al., 2001***) libraries, with data analysis and plotting using matplotlib (***Hunter, 2007***) and pandas (***McKinney, 2010***). The details of the physiological model itself are given below.

Biologically feasible simulation of stochastic tuning

To provide a suitable mechanistic model for stochastic tuning, we developed a discrete-time model tracking the temporal evolution of transcription rates $r_{i,t}$ (continuous, changed in response to

random fluctuations and potentially tuning input), copy number of each transcript per cell x_i, and copy number of each protein per cell p_i, considered separately for each gene i.

Transcriptional regulation lies at the center of our consideration for fitness-directed tuning. In the physiological model, there is a time-dependent probability $r_{i,t}$ for a single transcript to be generated from gene i at each timestep; the probabilities $r_{i,t}$ are updated in response to changing fitness as described below. In addition, each copy of the transcript present in the cell has a fixed probability d_i of being degraded at each timestep. The net change at each timestep t in the transcript level x_i for each gene i is thus given by

$$x_{i,t} \sim x_{i,t-1} - \text{binom}(x_{i,t-1}, d_i) + \text{bern}(r_{i,t-1})$$

Here binom/bern are binomial and Bernoulli random variables, respectively. Terms using binomial distributions allow a uniform probability for each present copy of a protein or transcript to be degraded or translated, whereas the Bernoulli term captures the probability of a transcript arising from each gene in a single timestep. We used a timestep of 1 s for all simulations described here.

Protein production in our physiological model arises from similar principles. At each timestep, each copy of a transcript from gene i has a fixed gene-dependent probability l_i of being translated to produce a single copy of the corresponding protein. In addition, each copy of that protein already present in the cell has a gene-dependent probability e_i of being degraded. Thus, the net rate of change in the protein copy number p_i at each time t is governed by the equation

$$p_{i,t} \sim p_{i,t-1} - \text{binom}(p_{i,t-1}, e_i) + \text{binom}(x_{i,t-1}, l_i)$$

The fixed, gene-specific parameters d_i, e_i, and l_i were drawn from distributions that are themselves fits to appropriate experimental data; we then modified the fitted parameters to yield distributions that are contained within the physiological distributions, while excluding the extreme ends of the available range. The parameters used for the physiological rate distributions are summarized below:

Transcription rates (used to initialize the transcription rate distribution, and separately to set the target transcription rate distribution): Transcripts per hour are gamma distributed with shape = 5 and rate = 2 (obtained by fitting data from (***Holstege et al., 1998***) and excluding extreme values)

Transcript degradation rates d_i: Half lives in minutes have a gamma distribution with shape = 12.0 and rate = 0.75 (obtained by fitting data from (***Holstege et al., 1998***) and then modifying to exclude extreme values).

Protein degradation rates e_i: Half lives in hours have a scaled t distribution with mean = 1, sigma = 0.382, and 80 degrees of freedom (fit based on data from (***Christiano et al., 2014***), but modified to exclude long half-lives, consistent with the induction of autophagy in stressed cells (***Cebollero and Reggiori, 2009***)).

Protein synthesis rates l_i: log2 synthesis rates per transcript have a scaled t distribution with mean=-5, sigma = 0.5, and 80 degrees of freedom (in units of s^{-1}); based on protein abundance data from (***Kulak et al., 2014***) combined with the other parameters defined above, and modified to exclude extreme values).

As described in the main text, our model permits two classes of 'marks' (representing histone modifications) that alter transcription rates: tuning marks (T), which change in level on the basis of recent changes in fitness and the current tuning mark state at each gene, and stabilizing marks (S), which change in abundance based on the tuning mark levels at each promoter. The number of each mark type at each promoter may be positive or negative, reflecting the possibility of distinct activating (+) or repressing (-) chromatin modifications.

The rate of change in the tuning marks proceeds according to the following principles. At each timestep, marks may be added or removed on the basis of recent changes in fitness; each mark may decay with a fixed probability; and marks may be added or removed in an undirected manner due to random drift. Referring to the number of tuning marks at a particular gene i as μ_i, the change in tuning marks at each timestep due to the tuning contribution alone is given by

$$\Delta\mu_{i,\text{tuning}} \sim \text{sgn}(\Delta F_t) * \text{sgn}(\mu_i) * \text{randint}(1,5) * \text{bern}(p_{\text{tunestep}})$$

Here sgn(x) is one if x is positive, -1 if x is negative, and 0 if x is zero. Δ F indicates the difference in mean fitness between the previous n_{window} steps and the nonoverlapping block of n_{window} steps before that; thus, sgn(ΔF_t) will be positive if the cells are becoming healthier, and negative if the cells are becoming less healthy. The fitness itself, F_t, is calculated as the Euclidean distance between the observed vector of protein levels $\boldsymbol{p_t}$ at a particular timestep, and the median observed in the last quarter of a long (10 times the normal simulation length) trajectory where all transcription rates are

fixed at their target values (note that oscillation still occurs, even in this case of known-correct transcription rates, due to the inherent randomness in transcript and protein production and degradation). In the context of our model, ΔF represents the direction of change in global cellular health, and $p_{tunestep}$ indicates the probability that tuning marks will be added/removed at a particular timestep. The combination of signs of the change in fitness (ΔF) and marks (μ_i) ensures that if the fitness is increasing and a given promoter has a positive number of tuning marks, the number of tuning marks at that promoter will increase further, whereas if the fitness was decreasing, the number of tuning marks will be decreased. The inverse directions apply for promoters with negative levels of T marks. Note that for control simulations where the effects of tuning are removed, the sign of the fitness-dependent term above is instead taken to be random.

The removal and random drift of tuning marks are governed by the equations

$\Delta\mu_{i,removal} \sim -1 * sgn(\mu_i) * binom(\mu_i, p_{decay})$

and

$\Delta\mu_{i,random} \sim (1–2*bern(0.5)) * bern(p_{random})$ respectively. The first equation here indicates that each individual mark may be removed with probability p_{decay} at each timestep, and in addition, the second equation dictates that each promoter may have a single mark of random sign added at each timestep, with probability p_{random}. The overall equation for the change in tuning marks at promoter *i* at each timestep is thus given by the sum of the terms above:

$\Delta\mu_{i,t} \sim \Delta F_t * sgn(\mu_{i,t-1}) * randint(1,5) * bern(p_{tunestep}) - sgn(\mu_i) * binom(\mu_i, p_{decay}) + (1–2*bern(0.5)) * bern(p_{random})$

The stabilizing marks (S), in contrast, do not vary directly in response to fitness, but rather, at each timestep may be added or removed from each promoter depending on its current state of T marks (see ***Figure 9B***): if the promoter has a high transcription rate due to high T levels, the net S count will be increased (with a probability at each timestep proportional to the current magnitude of the T level), and if the promoter has low T levels, the net S count is decreased. The effect of the stabilizing marks is to slowly shift the baseline transcription rate of genes over time. The change in number of S marks ν_i at gene *i* at each timestep is given by:

$\Delta\nu_{i,t} \sim sgn(\mu_{i,t-1}) * bern(abs(\mu_{i,t-1}) * p_{s\_mark} / \mu_{max})$

Here p_{s_mark} is a probability of changing S marks at each time step, and μ_{max} the maximum number of T marks allowed at a given promoter, whether positive (activating) or negative (repressive).

Every gene in the model is taken to have a baseline transcription rate, $r_{i,0}$, drawn from the physiological distributions defined above. The time-dependent instantaneous transcription rate of a given gene, $r_{i,t}$, is then calculated from the number of tuning marks (μ_i) and stabilizing marks (ν_i). The effects of tuning and stabilizing marks in the model are multiplicative, such that the transcription rate r_i at gene *i* with μ_i tuning marks and ν_i stabilizing marks is given by

$r_{i,t} = r_{i,0} * \alpha * exp(\beta)$; where $\alpha = 2*((\mu_{i,t} / \mu_{max})+1)$ and $\beta = m_S * \nu_{i,t}$

Here m_S represents the magnitude of the effects of a single S mark, and the number of T marks is constrained to the interval $[-\mu_{max}, \mu_{max}]$. The various fixed model parameters (e.g., m_S, p_{decay}, etc.) were chosen to be physiologically plausible while supporting tuning. The values of these parameters used in ***Figure 9C and E*** are taken as a baseline and shown in ***Supplementary file 8***; note, however, that as shown in ***Figure 9D***, the performance of the model is robust to changes in those parameters.

A python implementation of the model, along with sample inputs corresponding to the simulations described here, are included as ***Source code 3***.

Fluorescence tracking of sorted populations

In order to measure the mixing times under different stress conditions, synprom-URA3-mRuby/synprom-DHFR-GFP cells were grown overnight in SC +glu media. The next morning, the cells were back-diluted 1:100 into fresh, prewarmed low fluorescence SC +glu or ura-/6AU10. The cells in ura-/6AU10 media were kept in a 30˚C incubated shaker for 24 hr before sorting, whereas the cells in the complete media were sorted after four hours of growth at 30˚C. The cells were sorted based on their mRuby fluorescence level into three populations of the top 20%, bottom 20%, and the complete distribution (mock-sorted) of cells. In order to minimize the effects of both autofluorescence and size-fluorescence correlations, the cells (including those in the mock-sorted population) were tightly gated on FSC-A levels. The sorted cells were kept on ice until they were spun down and transferred to pre-warmed media identical to that in which they had previously been incubated (that is, cells from complete media to complete media and cells from ura-/6AU to fresh ura-/6AU). The cells were

incubated at 30°C thereafter. A sample of each population was analyzed using flow cytometry at different time intervals, with T = 0 being the time that the fresh media was added to the samples. The last time point for the cells in SC +glu media was 630 min, and for the ura-/6AU cells was 6660 min.

We calculated the distribution of mRuby fluorescence values for each sample at each time point by smoothing the observed values using a kernel density estimator. We then measured the pairwise mRuby fluorescence distribution overlap of the top 20%, bottom 20% and the complete distribution at each time point for each growth condition. The distribution overlap was calculated by numerically integrating the area under the (normalized) kernel density distribution estimates of both populations being compared.

An increasing overlap relative to t = 0 signifies the amount that the two populations have moved towards each other, and therefore the higher the overlap, the more mixed the two populations have become. Therefore, we calculated $f(t) = \frac{(\max(x) - x(t))}{x_{t=0}}$, where x is the overlap between the two distributions and max(x) is the maximum observed overlap. *f*(t) can be modeled as an exponential decay process according to:

$$f(t) = ae^{-\frac{t}{\tau}}$$

where τ provides a timescale for the mixing time (in particular, τ ln(2) is the half-life of the decay process). We used nonlinear curve fitting in Matlab to estimate the values of the parameters in the above equation for cells grown under each of the physiological conditions described above and report the estimated half-lives to give insight into the mixing times active in the populations studied here.

Fluorescence microscopy time courses on immobilized cells

The images shown and analyzed in ***Figure 6***, ***Figure 7***, and panel A of ***Figure 7—figure supplement 2*** were obtained on a Zeiss Axio Observer Z1, using a 40x objective lens. P_{HSP12}-URA3-mRuby/P_{ADH1}-DHFR-GFP cells were grown overnight in SC +glu liquid media, and then back-diluted 100x into SC/loflo + glu media and grown four additional hours with shaking at 30°C. Cells were spun down, and then incubated in ura-/loflo/6AU5 liquid for 12–13 hr. The cells were then pipetted onto the prepared slides. In order to prepare slides, we added 200 µL of ura-/loflo/6AU5 media containing 1% agar to each well of a two-well slide. Using a 22 µm coverslip, the surface of the media in the wells containing the solid media was flattened. After adding the cells on to the wells, we allowed extra media to be absorbed and then added a cover slip on top. The cells were imaged on DIC, GFP, brightfield, and mRuby channels; snapshots were taken once every 30 min for approximately 24 hr.

The additional imaging time series analyzed in panels B-D of ***Figure 7—figure supplement 2*** were obtained for P_{HSP12}-URA3-mRuby/P_{ADH1}-DHFR-GFP cells immobilized to thin-bottomed growth chambers and grown in ura-/6AU5 media. To prepare the slides, cells were grown overnight in SC +glu liquid media, and then back-diluted 100x into SC/loflo + glu media and grown four additional hours with shaking at 30°C. During that incubation, a coverslip/incubation chamber (Nunc) was treated for five minutes with poly-D-lysine solution (MPI Biomedical), washed three times with sterile deionized water, and then allowed to dry.

After the pregrowth period, cells were diluted 10x into additional prewarmed SC/loflo + glu, and then pipetted onto the poly-D-lysine treated cover slip and allowed to settle for 30 min at room temperature. The media was removed, and non-adherent cells were washed away with two 1 mL rinses of sterile deionized water. The cells were then covered with 2 mL of ura-/loflo/6AU5 media, and then placed in a preheated microscopy incubation chamber (OKO) at 30°C and 90% relative humidity. Cells were imaged on DIC, GFP, and mRuby channels; snapshots taken once every 30 min for 24 hr on a Nikon Eclipse Ti microscope using a 20x objective.

For comparative visualization purposes (***Figure 7A–B***), the DIC or brightfield channel of each image was rescaled using the ImageMagick 'normalize' operator, and the fluorescence channels were normalized by subtracting the minimum pixel value within a given field of view, and then subjecting the remaining data to a median filter over a 5 × 5 pixel window. The fluorescence channels were then stacked on the DIC or brightfield to generate the images shown. Un-normalized data were used for all quantitative analysis.

For the quantitative analysis in ***Figure 7C–E*** and ***Figure 7—figure supplement 2***, segmentation and lineage tracking were performed manually to identify cell division events and define cell interiors at the plotted timepoints. The fluorescence of each cell for each channel was then taken to be the average value of all pixels within the defined cell interior, with the mode value of all pixels in a defined window around the cell subtracted as background. For the purpose of classifying cells based on their division state, a cell was classified as 'dividing' if it gave rise to a daughter cell before the next analyzed snapshot. Timepoints prior to three hours were excluded from quantitative analysis of dividing vs. nondividing cells for the populations pregrown in SC/loflo + glu, as a large fraction of cells in all of our microscopy experiments did undergo a single division before arresting, likely using residual nutrients from their previous growth in complete media.

Acknowledgements

We thank members of the Tavazoie laboratory and Sohail Tavazoie for helpful discussions and feedback on the manuscript, and Panos Oikonomou for crucial input on the use of sorted cells to track mixing times. We are also grateful to Kaushik Ragunathan for additional helpful discussion and access to microscopy equipment. PLF was supported by a K99 Award (K99/R00-GM097033) from NIGMS. JY was supported by the NIH MSTP program at Columbia University Medical School. ST was supported by grants from NIH/NIAID (R01-AI077562) and the NIH Director's Pioneer Award (DP1-ES022578).

Additional information

Funding

Funder	Grant reference number	Author
National Institute of General Medical Sciences	K99 (GM097033-01A1)	Peter L Freddolino
National Institutes of Health	MSTP	Jamie Yang
NIH Office of the Director	8DP1ES022578	Saeed Tavazoie
National Institute of Allergy and Infectious Diseases	R01-AI077562	Saeed Tavazoie

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Peter L Freddolino, Conceptualization, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Jamie Yang, Data curation, Validation, Methodology, Writing—review and editing; Amir Momen-Roknabadi, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Saeed Tavazoie, Conceptualization, Software, Formal analysis, Supervision, Funding acquisition, Investigation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing

Author ORCIDs

Peter L Freddolino http://orcid.org/0000-0002-5821-4226
Saeed Tavazoie http://orcid.org/0000-0003-2183-4162

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.31867.051
Author response https://doi.org/10.7554/eLife.31867.052

Additional files

Supplementary files

- Source code 1. Custom R and python code for analysis of flow cytometry data.

DOI: https://doi.org/10.7554/eLife.31867.020

- Source code 2. Octave implementation of the simple stochastic tuning model used in ***Figures 1***, ***2*** and ***5***. See the accompanying README file for documentation.

DOI: https://doi.org/10.7554/eLife.31867.021

- Source code 3. Python implementation of the physiological stochastic tuning model used in ***Figure 9***. See the accompanying README file for documentation.

DOI: https://doi.org/10.7554/eLife.31867.022

- Supplementary file 1. Complete sequence of synprom5-sam3, the promoter referred to as 'synprom' in the main text. Pseudorandomly generated sequence is shown on the first line, and SAM3-derived sequence on the second. The sequence has been perturbed to remove all recognizable transcription factor binding sites, as described in Materials and Methods.

DOI: https://doi.org/10.7554/eLife.31867.023

- Supplementary file 2. Sequences of other synthetic promoter components referenced in the text. Synthetic promoters consist of the combination of one pseudorandom 'synprom' sequence with either the SAM3 or ARF1 promoter proximal regions. All sequences have been perturbed to remove all recognizable transcription factor binding sites, as described in Experimental Procedures.

DOI: https://doi.org/10.7554/eLife.31867.024

- Supplementary file 3. Fitted parameters for distribution overlap half-lives from Figure S4. Shown are fitted values for the half-lives plus (in parentheses) the extent of a 95% confidence interval based on the model fit. All half-lives are given in minutes.

DOI: https://doi.org/10.7554/eLife.31867.025

- Supplementary file 4. Results of resequencing of tuned colonies and planktonic populations in the 25 kb vicinity of the URA3 and DHFR insertions. Numbers in parenthesis after mutation calls indicate the approximate fraction of the population containing the mutant allele. 'ID' is simply an identifier used to refer to each sample in the text.

DOI: https://doi.org/10.7554/eLife.31867.026

- Supplementary file 5. Colony counts for the extreme most highly fluorescent cells (top 0.5–1%) isolated from populations in which URA3-mRuby is driven by the specified promoter. Equal volumes of the sorted cells were plated in parallel on SC+glu and ura-/6AU15 plates, and then counted after 2–3 days (SC+glu) or 19–20 days (6AU). 'Baseline' refers to the fraction of cells expected to form colonies on 6AU15 plates in 19–20 days in unsorted populations (c.f. ***Figures 3–4*** of the main text).

DOI: https://doi.org/10.7554/eLife.31867.027

- Supplementary file 6. Codon optimized sequence of superfolder GFP used in all GFP constructs. Note that no start codon is included, as the construct is intended to be part of a fusion protein.

DOI: https://doi.org/10.7554/eLife.31867.028

- Supplementary file 7. Primer design for quantitative PCR experiments. End locations are given relative to the start codon of the gene in question.

DOI: https://doi.org/10.7554/eLife.31867.029

- Supplementary file 8. Baseline model parameters for the physiological tuning simulations described in ***Figure 9***.

DOI: https://doi.org/10.7554/eLife.31867.030

- Transparent reporting form

DOI: https://doi.org/10.7554/eLife.31867.031

Major datasets

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Freddolino PL, Yang J, Tavazoie S	2017	Saccharomyces cerevisiae genome sequencing after transcriptional tuning	https://trace.ddbj.nig.ac.jp/DRASearch/study?acc=SRP117724	Publicly available at the DNA Data Bank of Japan (accession no. SRA608529)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652631	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652631)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652632	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652632)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652633	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652633)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652634	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652634)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652635	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652635)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652636	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652636)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652637	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652637)
Freddolino PL, Yang J, Tavazoie S	2017	NGS data from Stochastic tuning of gene expression enables cellular adaptation in the absence of pre-existing regulatory circuitry	https://www.ncbi.nlm.nih.gov/sra/?term=SAMN07652638	Publicly available at the NCBI Sequence Read Archive (accession no: SAMN07652638)

References

Akada R, Hirosawa I, Kawahata M, Hoshida H, Nishizawa Y. 2002. Sets of integrating plasmids and gene disruption cassettes containing improved counter-selection markers designed for repeated use in budding yeast. *Yeast* **19**:393–402. DOI: https://doi.org/10.1002/yea.841, PMID: 11921088

Al Mamun AA, Lombardo MJ, Shee C, Lisewski AM, Gonzalez C, Lin D, Nehring RB, Saint-Ruf C, Gibson JL, Frisch RL, Lichtarge O, Hastings PJ, Rosenberg SM. 2012. Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science* **338**:1344–1348. DOI: https://doi.org/10.1126/science.1226683, PMID: 23224554

Albert V, Hall MN. 2015. mTOR signaling in cellular and organismal energetics. *Current Opinion in Cell Biology* **33**:55–66. DOI: https://doi.org/10.1016/j.ceb.2014.12.001, PMID: 25554914

Applebee MK, Herrgård MJ, Palsson BØ. 2008. Impact of individual mutations on increased fitness in adaptively evolved strains of Escherichia coli. *Journal of Bacteriology* **190**:5087–5094. DOI: https://doi.org/10.1128/JB.01976-07, PMID: 18487343

Álvarez-Errico D, Vento-Tormo R, Sieweke M, Ballestar E. 2015. Epigenetic control of myeloid cell differentiation, identity and function. *Nature Reviews Immunology* **15**:7–17. DOI: https://doi.org/10.1038/nri3777, PMID: 25534619

Barenholz U, Keren L, Segal E, Milo R. 2016. A Minimalistic Resource Allocation Model to Explain Ubiquitous Increase in Protein Expression with Growth Rate. *PLoS One* **11**:e0153344. DOI: https://doi.org/10.1371/journal.pone.0153344, PMID: 27073913

Basehoar AD, Zanton SJ, Pugh BF. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**:699–709. DOI: https://doi.org/10.1016/S0092-8674(04)00205-3, PMID: 15006352

Bitterman KJ, Anderson RM, Cohen HY, Latorre-Esteves M, Sinclair DA. 2002. Inhibition of silencing and accelerated aging by nicotinamide, a putative negative regulator of yeast sir2 and human SIRT1. *Journal of Biological Chemistry* **277**:45099–45107. DOI: https://doi.org/10.1074/jbc.M205670200, PMID: 12297502

Blake WJ, Balázsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, Cantor CR, Walt DR, Collins JJ. 2006. Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular Cell* **24**:853–865. DOI: https://doi.org/10.1016/j.molcel.2006.11.003, PMID: 17189188

Blount ZD, Barrick JE, Davidson CJ, Lenski RE. 2012. Genomic analysis of a key innovation in an experimental Escherichia coli population. *Nature* **489**:513–518. DOI: https://doi.org/10.1038/nature11514, PMID: 22992527

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114–2120. DOI: https://doi.org/10.1093/bioinformatics/btu170, PMID: 24695404

Bonito NA, Borley J, Wilhelm-Benartzi CS, Ghaem-Maghami S, Brown R. 2016. Epigenetic regulation of the homeobox gene MSX1 associates with platinum-resistant disease in high-grade serous epithelial ovarian cancer. *Clinical Cancer Research* **22**:3097–3104. DOI: https://doi.org/10.1158/1078-0432.CCR-15-1669, PMID: 26763252

Borley J, Brown R. 2015. Epigenetic mechanisms and therapeutic targets of chemotherapy resistance in epithelial ovarian cancer. *Annals of Medicine* **47**:359–369. DOI: https://doi.org/10.3109/07853890.2015.1043140, PMID: 26158617

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD. 1998. Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**:115–132. DOI: https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2, PMID: 9483801

Brauer MJ, Huttenhower C, Airoldi EM, Rosenstein R, Matese JC, Gresham D, Boer VM, Troyanskaya OG, Botstein D. 2008. Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Molecular Biology of the Cell* **19**:352–367. DOI: https://doi.org/10.1091/mbc.E07-08-0779, PMID: 17959824

Broach JR. 2012. Nutritional control of growth and development in yeast. *Genetics* **192**:73–105. DOI: https://doi.org/10.1534/genetics.111.135731, PMID: 22964838

Cauchy A. 1847. Methode generale pour la resolution des systemes d'equations simultanees. *Comptes Rendus Hebdomadaires Des Seances De l'Academie Des Sciences* **25**:536–538.

Cebollero E, Reggiori F. 2009. Regulation of autophagy in yeast Saccharomyces cerevisiae. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research* **1793**:1413–1421. DOI: https://doi.org/10.1016/j.bbamcr.2009.01.008, PMID: 19344676

Chen H, Fan M, Pfeffer LM, Laribee RN. 2012. The histone H3 lysine 56 acetylation pathway is regulated by target of rapamycin (TOR) signaling and functions directly in ribosomal RNA biogenesis. *Nucleic Acids Research* **40**:6534–6546. DOI: https://doi.org/10.1093/nar/gks345, PMID: 22553361

Christiano R, Nagaraj N, Fröhlich F, Walther TC. 2014. Global proteome turnover analyses of the Yeasts S. cerevisiae and S. pombe. *Cell Reports* **9**:1959–1965. DOI: https://doi.org/10.1016/j.celrep.2014.10.065, PMID: 25466257

Collart MA, Oliviero S. 2001. Preparation of Yeast RNA. In: *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc.

Conrad M, Schothorst J, Kankipati HN, Van Zeebroeck G, Rubio-Texeira M, Thevelein JM. 2014. Nutrient sensing and signaling in the yeast Saccharomyces cerevisiae. *FEMS Microbiology Reviews* **38**:254–299. DOI: https://doi.org/10.1111/1574-6976.12065, PMID: 24483210

Damkiær S, Yang L, Molin S, Jelsbak L. 2013. Evolutionary remodeling of global regulatory networks during long-term bacterial adaptation to human hosts. *PNAS* **110**:7766–7771. DOI: https://doi.org/10.1073/pnas.1221466110, PMID: 23610385

David L, Ben-Harosh Y, Stolovicki E, Moore LS, Nguyen M, Tamse R, Dean J, Mancera E, Steinmetz LM, Braun E. 2013. Multiple genomic changes associated with reorganization of gene regulation and adaptation in yeast. *Molecular Biology and Evolution* **30**:1514–1526. DOI: https://doi.org/10.1093/molbev/mst071, PMID: 23589456

David L, Stolovicki E, Haziz E, Braun E. 2010. Inherited adaptation of genome-rewired cells in response to a challenging environment. *HFSP Journal* **4**:131–141. DOI: https://doi.org/10.2976/1.3353782, PMID: 20811567

DeMarini DJ, Carlin EM, Livi GP. 2001. Constitutive promoter modules for PCR-based gene modification in Saccharomyces cerevisiae. *Yeast* **18**:723–728. DOI: https://doi.org/10.1002/yea.721, PMID: 11378899

Eaton JW, Bateman D, Hauberg S. 2009. *GNU Octave Version 3.0.1 Manual: A High-Level Interactive Language for Numerical Computations*. CreateSpace Independent Publishing Platform.

Ellis B, Gentleman R, Hahne F, Meur NL, Sarkar D. 2006. Visualization for flow cytometry. *flowViz*. 1.30.0.

Ellis B, Haaland P, Hahne F, Meur NL, Gopalakrishnan N, Spidlen J. 2009. flowCore: Basic structures for flow cytometry data. *flowCore*. 1.32.1.

Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* **297**:1183–1186. DOI: https://doi.org/10.1126/science.1070919, PMID: 12183631

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K,

Deutschbauer A, et al. 2002. Functional profiling of the Saccharomyces cerevisiae genome. *Nature* **418**:387–391. DOI: https://doi.org/10.1038/nature00935, PMID: 12140549

Gietz RD, Woods RA. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in Enzymology* **350**:87–96. DOI: https://doi.org/10.1016/S0076-6879(02)50957-5, PMID: 12073338

González A, Hall MN. 2017. Nutrient sensing and TOR signaling in yeast and mammals. *The EMBO Journal* **36**: 397–408. DOI: https://doi.org/10.15252/embj.201696010, PMID: 28096180

Goodarzi H, Bennett BD, Amini S, Reaves ML, Hottes AK, Rabinowitz JD, Tavazoie S. 2010. Regulatory and metabolic rewiring during laboratory evolution of ethanol tolerance in E. coli. *Molecular Systems Biology* **6**:378. DOI: https://doi.org/10.1038/msb.2010.33, PMID: 20531407

Guan Q, Haroon S, Bravo DG, Will JL, Gasch AP. 2012. Cellular memory of acquired stress resistance in Saccharomyces cerevisiae. *Genetics* **192**:495–505. DOI: https://doi.org/10.1534/genetics.112.143016, PMID: 22851651

Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**:717–728. DOI: https://doi.org/10.1016/S0092-8674(00)81641-4, PMID: 9845373

Hottes AK, Freddolino PL, Khare A, Donnell ZN, Liu JC, Tavazoie S. 2013. Bacterial adaptation through loss of function. *PLoS Genetics* **9**:e1003617. DOI: https://doi.org/10.1371/journal.pgen.1003617, PMID: 23874220

Hunter JD. 2007. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**:90–95. DOI: https://doi.org/10.1109/MCSE.2007.55

Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**:318–356. DOI: https://doi.org/10.1016/S0022-2836(61)80072-7, PMID: 13718526

Jeffreys H. 1961. *Theory of Probability*. Oxford University Press.

Jones E, Oliphant T, Peterson P. 2001. *Scipy: Open Source Scientific Tools for Python*.

Kaiser C, Michaelis S, Mitchell A. 1994. *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Katada S, Imhof A, Sassone-Corsi P. 2012. Connecting threads: epigenetics and metabolism. *Cell* **148**:24–28. DOI: https://doi.org/10.1016/j.cell.2012.01.001, PMID: 22265398

Keren L, Zackay O, Lotan-Pompan M, Barenholz U, Dekel E, Sasson V, Aidelberg G, Bren A, Zeevi D, Weinberger A, Alon U, Milo R, Segal E. 2013. Promoters maintain their relative activity levels under different growth conditions. *Molecular Systems Biology* **9**:701. DOI: https://doi.org/10.1038/msb.2013.59, PMID: 24169404

Klumpp S, Hwa T. 2014. Bacterial growth: global effects on gene expression, growth feedback and proteome partition. *Current Opinion in Biotechnology* **28**:96–102. DOI: https://doi.org/10.1016/j.copbio.2014.01.001, PMID: 24495512

Klumpp S, Zhang Z, Hwa T. 2009. Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**:1366–1375. DOI: https://doi.org/10.1016/j.cell.2009.12.001, PMID: 20064380

Kredel S, Oswald F, Nienhaus K, Deuschle K, Röcker C, Wolff M, Heilker R, Nienhaus GU, Wiedenmann J. 2009. mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. *PLoS One* **4**:e4391. DOI: https://doi.org/10.1371/journal.pone.0004391, PMID: 19194514

Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M, Minimal MM. 2014. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods* **11**:319–324. DOI: https://doi.org/10.1038/nmeth.2834, PMID: 24487582

Kurdistani SK. 2014. Chromatin: a capacitor of acetate for integrated regulation of gene expression and cell physiology. *Current Opinion in Genetics & Development* **26**:53–58. DOI: https://doi.org/10.1016/j.gde.2014.06.002, PMID: 25016437

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**:R25. DOI: https://doi.org/10.1186/gb-2009-10-3-r25, PMID: 19261174

Lartigue C, Vashee S, Algire MA, Chuang RY, Benders GA, Ma L, Noskov VN, Denisova EA, Gibson DG, Assad-Garcia N, Alperovich N, Thomas DW, Merryman C, Hutchison CA, Smith HO, Venter JC, Glass JI. 2009. Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325**:1693–1696. DOI: https://doi.org/10.1126/science.1173759, PMID: 19696314

Li B, Carey M, Workman JL. 2007. The role of chromatin during transcription. *Cell* **128**:707–719. DOI: https://doi.org/10.1016/j.cell.2007.01.015, PMID: 17320508

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079. DOI: https://doi.org/10.1093/bioinformatics/btp352, PMID: 19505943

Li S, Yue D, Chen X, Wang L, Li J, Ping Y, Gao Q, Wang D, Zhang T, Li F, Yang L, Huang L, Zhang Y. 2015. Epigenetic regulation of CD271, a potential cancer stem cell marker associated with chemoresistance and metastatic capacity. *Oncology Reports* **33**:425–432. DOI: https://doi.org/10.3892/or.2014.3569, PMID: 25351876

Lin SJ, Defossez PA, Guarente L. 2000. Requirement of NAD and SIR2 for life-span extension by calorie restriction in Saccharomyces cerevisiae. *Science* **289**:2126–2128. DOI: https://doi.org/10.1126/science.289.5487.2126, PMID: 11000115

Luka Z, Mudd SH, Wagner C. 2009. Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. *Journal of Biological Chemistry* **284**:22507–22511. DOI: https://doi.org/10.1074/jbc.R109.019273, PMID: 19483083

Lv JF, Hu L, Zhuo W, Zhang CM, Zhou HH, Fan L. 2016. Epigenetic alternations and cancer chemotherapy response. *Cancer Chemotherapy and Pharmacology* **77**:673–684. DOI: https://doi.org/10.1007/s00280-015-2951-0, PMID: 26707728

Macnab RM, Koshland DE. 1972. The gradient-sensing mechanism in bacterial chemotaxis. *PNAS* **69**:2509–2512. DOI: https://doi.org/10.1073/pnas.69.9.2509, PMID: 4560688

Martin M. 2014. *Algorithms and Tools for the Analysis of High Throughput DNA Sequencing Data*. Technische Universitat Dortmund.

McKinney W. 2010. Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*.

Meissner A. 2010. Epigenetic modifications in pluripotent and differentiated cells. *Nature Biotechnology* **28**:1079–1088. DOI: https://doi.org/10.1038/nbt.1684, PMID: 20944600

Mitra D, Parnell EJ, Landon JW, Yu Y, Stillman DJ. 2006. SWI/SNF binding to the HO promoter requires histone acetylation and stimulates TATA-binding protein recruitment. *Molecular and Cellular Biology* **26**:4095–4110. DOI: https://doi.org/10.1128/MCB.01849-05, PMID: 16705163

Moore LS, Wei W, Stolovicki E, Benbenishty T, Wilkening S, Steinmetz LM, Braun E, David L. 2014. Induced mutations in yeast cell populations adapting to an unforeseen challenge. *PLoS One* **9**:e111133. DOI: https://doi.org/10.1371/journal.pone.0111133, PMID: 25340744

Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. 2006. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**:840–846. DOI: https://doi.org/10.1038/nature04785, PMID: 16699522

Perez-Plasencia C, Duenas-Gonzalez A. 2006. Can the state of cancer chemotherapy resistance be reverted by epigenetic therapy? *Molecular Cancer* **5**:27. DOI: https://doi.org/10.1186/1476-4598-5-27, PMID: 16831224

Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. 2006. Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology* **24**:79–88. DOI: https://doi.org/10.1038/nbt1172, PMID: 16369541

Philippe N, Crozat E, Lenski RE, Schneider D. 2007. Evolution of global regulatory networks during a long-term experiment with Escherichia coli. *BioEssays* **29**:846–860. DOI: https://doi.org/10.1002/bies.20629, PMID: 17691099

Plummer M. 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria. https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf.

Potrykus K, Murphy H, Philippe N, Cashel M. 2011. ppGpp is the major source of growth rate control in E. coli. *Environmental Microbiology* **13**:563–575. DOI: https://doi.org/10.1111/j.1462-2920.2010.02357.x, PMID: 20946586

Raj A, van Oudenaarden A. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**:216–226. DOI: https://doi.org/10.1016/j.cell.2008.09.050, PMID: 18957198

Rando OJ, Winston F. 2012. Chromatin and transcription in yeast. *Genetics* **190**:351–387. DOI: https://doi.org/10.1534/genetics.111.132266, PMID: 22345607

Raser JM, O'Shea EK. 2004. Control of stochasticity in eukaryotic gene expression. *Science* **304**:1811–1814. DOI: https://doi.org/10.1126/science.1098641, PMID: 15166317

Raser JM, O'Shea EK. 2005. Noise in gene expression: origins, consequences, and control. *Science* **309**:2010–2013. DOI: https://doi.org/10.1126/science.1105891, PMID: 16179466

Rodríguez-Verdugo A, Tenaillon O, Gaut BS. 2016. First-step mutations during adaptation restore the expression of hundreds of genes. *Molecular Biology and Evolution* **33**:25–39. DOI: https://doi.org/10.1093/molbev/msv228, PMID: 26500250

Russo P, Sherman F. 1989. Transcription terminates near the poly(A) site in the CYC1 gene of the yeast Saccharomyces cerevisiae. *PNAS* **86**:8348–8352. DOI: https://doi.org/10.1073/pnas.86.21.8348, PMID: 2554310

Sanchez A, Golding I. 2013. Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**:1188–1193. DOI: https://doi.org/10.1126/science.1242975, PMID: 24311680

Saxton RA, Sabatini DM. 2017. mTOR Signaling in Growth, Metabolism, and Disease. *Cell* **168**:960–976. DOI: https://doi.org/10.1016/j.cell.2017.02.004, PMID: 28283069

Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, Beqiri M, Sproesser K, Brafford PA, Xiao M, Eggan E, Anastopoulos IN, Vargas-Garcia CA, Singh A, Nathanson KL, Herlyn M, Raj A. 2017. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**:431–435. DOI: https://doi.org/10.1038/nature22794, PMID: 28607484

Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Liron Y, Rosenfeld N, Danon T, Perzov N, Alon U. 2006. Variability and memory of protein levels in human cells. *Nature* **444**:643–646. DOI: https://doi.org/10.1038/nature05316, PMID: 17122776

Simonsen CC, Levinson AD. 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *PNAS* **80**:2495–2499. DOI: https://doi.org/10.1073/pnas.80.9.2495, PMID: 6573667

Sourjik V, Wingreen NS. 2012. Responding to chemical gradients: bacterial chemotaxis. *Current Opinion in Cell Biology* **24**:262–268. DOI: https://doi.org/10.1016/j.ceb.2011.11.008, PMID: 22169400

Spivak AT, Stormo GD. 2012. ScerTF: a comprehensive database of benchmarked position weight matrices for Saccharomyces species. *Nucleic Acids Research* **40**:D162–D168. DOI: https://doi.org/10.1093/nar/gkr1180, PMID: 22140105

Stasevich TJ, Hayashi-Takanaka Y, Sato Y, Maehara K, Ohkawa Y, Sakata-Sogawa K, Tokunaga M, Nagase T, Nozaki N, McNally JG, Kimura H. 2014. Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature* **516**:272–275. DOI: https://doi.org/10.1038/nature13714, PMID: 25252976

Stern S, Dror T, Stolovicki E, Brenner N, Braun E. 2007. Genome-wide transcriptional plasticity underlies cellular adaptation to novel challenge. *Molecular Systems Biology* **3**:106. DOI: https://doi.org/10.1038/msb4100147, PMID: 17453047

Stolovicki E, Dror T, Brenner N, Braun E. 2006. Synthetic gene recruitment reveals adaptive reprogramming of gene regulation in yeast. *Genetics* **173**:75–85. DOI: https://doi.org/10.1534/genetics.106.055442, PMID: 16510783

Subramani S, Mulligan R, Berg P. 1981. Expression of the mouse dihydrofolate reductase complementary deoxyribonucleic acid in simian virus 40 vectors. *Molecular and Cellular Biology* **1**:854–864. DOI: https://doi.org/10.1128/MCB.1.9.854, PMID: 9279398

Svd W, Colbert SC, Varoquaux G. 2011. The NumPy Array: a structure for efficient numerical computation. *Computing in Science & Engineering* **13**:22–30. DOI: https://doi.org/10.1109/MCSE.2011.37

Tagkopoulos I, Liu YC, Tavazoie S. 2008. Predictive behavior within microbial genetic networks. *Science* **320**:1313–1317. DOI: https://doi.org/10.1126/science.1154456, PMID: 18467556

Tanner KG, Trievel RC, Kuo MH, Howard RM, Berger SL, Allis CD, Marmorstein R, Denu JM. 1999. Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. *Journal of Biological Chemistry* **274**:18157–18160. DOI: https://doi.org/10.1074/jbc.274.26.18157, PMID: 10373413

Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS. 2012. The molecular diversity of adaptive convergence. *Science* **335**:457–461. DOI: https://doi.org/10.1126/science.1212986, PMID: 22282810

Thaminy S, Newcomb B, Kim J, Gatbonton T, Foss E, Simon J, Bedalov A. 2007. Hst3 is regulated by Mec1-dependent proteolysis and controls the S phase checkpoint and sister chromatid cohesion by deacetylating histone H3 at lysine 56. *Journal of Biological Chemistry* **282**:37805–37814. DOI: https://doi.org/10.1074/jbc.M706384200, PMID: 17977840

Thieffry D, Huerta AM, Pérez-Rueda E, Collado-Vides J. 1998. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in Escherichia coli. *BioEssays* **20**:433–440. DOI: https://doi.org/10.1002/(SICI)1521-1878(199805)20:5<433::AID-BIES10>3.0.CO;2-2, PMID: 9670816

Tirosh I, Barkai N, Verstrepen KJ. 2009. Promoter architecture and the evolvability of gene expression. *Journal of Biology* **8**:95. DOI: https://doi.org/10.1186/jbiol204, PMID: 20017897

Tirosh I, Weinberger A, Bezalel D, Kaganovich M, Barkai N. 2008. On the relation between promoter divergence and gene expression evolution. *Molecular Systems Biology* **4**:159. DOI: https://doi.org/10.1038/msb4100198, PMID: 18197176

Tirosh I, Weinberger A, Carmi M, Barkai N. 2006. A genetic signature of interspecies variations in gene expression. *Nature Genetics* **38**:830–834. DOI: https://doi.org/10.1038/ng1819, PMID: 16783381

Van Hofwegen DJ, Hovde CJ, Minnich SA. 2016. Rapid evolution of citrate utilization by Escherichia coli by direct selection requires citT and dctA. *Journal of Bacteriology* **198**:1022–1034. DOI: https://doi.org/10.1128/JB.00831-15, PMID: 26833416

Workman JJ, Chen H, Laribee RN. 2016. Saccharomyces cerevisiae TORC1 controls histone acetylation by signaling through the Sit4/PP6 phosphatase to regulate sirtuin deacetylase nuclear accumulation. *Genetics* **203**:1733–1746. DOI: https://doi.org/10.1534/genetics.116.188458, PMID: 27343235

Wu JN, Roberts CW. 2013. ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discovery* **3**:35–43. DOI: https://doi.org/10.1158/2159-8290.CD-12-0361, PMID: 23208470

Zhou BO, Zhou JQ. 2011. Recent transcription-induced histone H3 lysine 4 (H3K4) methylation inhibits gene reactivation. *Journal of Biological Chemistry* **286**:34770–34776. DOI: https://doi.org/10.1074/jbc.M111.273128, PMID: 21849496

Ziller MJ, Edri R, Yaffe Y, Donaghey J, Pop R, Mallard W, Issner R, Gifford CA, Goren A, Xing J, Gu H, Cachiarelli D, Tsankov A, Epstein C, Rinn JR, Mikkelsen TS, Kohlbacher O, Gnirke A, Bernstein BE, Elkabetz Y, et al. 2015. Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* **518**:355–359. DOI: https://doi.org/10.1038/nature13990, PMID: 25533951