The cost of gene expression underlies a fitness trade-off in yeast

Gregory I. Langa,1, Andrew W. Murrayb,1,2, and David Botsteinb,1,2

aLewis-Sigler Institute for Integrative Genomics and the Department of Molecular Biology, Princeton University, Princeton, NJ 08544; and bFAS Center for Systems Biology and the Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

Contributed by David Botstein, February 12, 2009 (sent for review December 10, 2008)

Natural selection optimizes an organism’s genotype within the context of its environment. Adaptations to one environment can decrease fitness in another, revealing evolutionary trade-offs. Here, we show that the cost of gene expression underlies a trade-off between growth rate and mating efficiency in the yeast Saccharomyces cerevisiae. During asexual growth, mutations that eliminate the ability to mate provide an ~2% per-generation growth-rate advantage. Some strains, including most laboratory strains, carry an allele of GPA1 (an upstream component of the mating pathway) that increases mating efficiency by ~30% per round of mating at the cost of an ~1% per-generation growth-rate disadvantage. In addition to demonstrating a trade-off between growth rate and mating efficiency, our results illustrate differences in the selective pressures defining fitness in the laboratory versus the natural environment and show that selection, acting on the cost of gene expression, can optimize expression levels and promote gene loss.

Abbreviations: GPA1, galactose permease 1; MAP, mitogen-activated protein; MATA, mother (male) type; MATα, father (female) type.

A frequent observation in evolution is that traits not maintained by selection will be lost—this holds true at the morphological level and at the genetic level. Examples of gene loss include the loss of olfactory receptors in primates (1), the loss of pigmentmentation and vision in Astyanax cavefish (2), the loss of the galactose utilization pathway in yeast (3), and the degeneration of genes involved in carbon utilization during domestication of Streptococcus thermophilus (4). Such regressive evolution also occurs in laboratory populations; reduction in catabolic expression and a positive growth-rate advantage (Fig. 1A). It appears from these data that at least some sterile mutants have a clear growth-rate advantage over wild type. To determine whether all sterile strains have a similar advantage, and to determine the basis for any growth-rate advantage in the sterile strains, we used a combination of 4 methods: Phenotypic characterization of the spontaneous αFR mutants, growth-rate assays on targeted gene deletions within the mating pathway, mapping of the mutations in the most fit sterile strains, and expression analysis on αFR strains both with and without a growth-rate advantage.

The yeast mating pathway is one of the best studied mitogen-activated protein (MAP) kinase cascades (12). At the beginning of the pathway is a pheromone receptor (Ste2 in MATA or Ste3 in MATα) that binds the cognate mating pheromone. Receptor stimulation activates a heterotrimeric G protein (consisting of Gpa1, Ste18, and Ste4), which in turn, activates MAP kinase cascade (consisting of the MAP kinase kinase, Ste11, the MAP kinase kinase, Ste7, the MAP kinases Fus3 and Kss1, and

Sterility increases growth rate by eliminating unnecessary gene expression. We tested the hypothesis that sterile strains generally have a growth-rate advantage by isolating sterile mutants and testing their fitness. Haploid, a-mating type (MATα) cells, arrest in G1 when exposed to the mating pheromone, alpha-factor (αF), and thus cannot form colonies on media containing αF. We initiated, from a single colony of haploid MATα cells, a large number of parallel cultures that were plated onto either rich media or rich media containing αF. On rich media, the vast majority of cells form colonies, but on αF, only the small fraction of cells that have acquired mutations in pheromone-induced signaling can form colonies. From each culture, we randomly chose a single alpha-factor resistant (αFR) or unselected colony and measured its relative growth rate by using a FACS-based competitive growth-rate assay that can detect growth-rate differences as small as 0.5%. The growth-rate coefficient is a measure of the growth-rate advantage over wild type. Fig. 1A shows the growth-rate coefficient (s) for 27 unselected clones and 45 αFR clones. As a control we measured the relative growth rates of 24 similarly selected mutants that were resistant to canavanine, a toxic arginine analog. In each case, several clones have a low growth rate (s < −1%), suggesting that these strains have become mitochondrial deficient or have acquired a deleterious mutation. Excluding clones with s < −1%, the growth-rate coefficients of the unselected clones follow a tight distribution (Fig. 1A, sα = 0.08% ± 0.35%) indistinguishable from the distribution of the canavanine-resistant mutants (Fig. 1A, sα = 0.36% ± 0.48%, P > 0.05, Wilcoxon rank sum test); however, the growth-rate coefficients of the αFR mutants show greater variation and a positive growth-rate advantage (Fig. 1A, sα = 1.48% ± 0.85%, P < 10−7, Wilcoxon).
transformed strains arrest after Ste4 overexpression indicating a
however, loss of the receptor (Ste2) or Far1 does not 

tatively, (Fig. 1B) (P = 2.8 x 10^-5, 5.8 x 10^-3, 2.2 x 10^-6, respectively, t test); however, deletion of the α receptor (Ste2) or Far1 does not (Fig. 1B) [P = 0.23, 0.03, respectively, t test (the value for far1Δ is not significant because we are testing 5 deletion mutants and thus require < 0.05/5 = 0.01 to regard a result as significant)]. This suggests that a growth-rate advantage exists for the subset of sterile strains that abolish basal signaling through the pathway (therefore eliminating basal expression of the mating genes), which depends on Ste4, Ste7, and Ste12, but not Ste2 or Far1 (14).

We identified the mutations in several αFR mutants from the higher end of the growth-rate distribution by hybridizing genomic DNA to microarrays that cover the entire yeast genome (tiling arrays) and characterized their effect on basal expression downstream of Ste12 by using gene expression microarrays. For simplicity the 45 spontaneous αFR mutants were numbered in order of growth-rate advantage, from αFR-1 (highest) to αFR-45 (lowest). We chose 7 αFR strains for this analysis: αFR-1, -2, -4, -7, -8, -17, and -20; our preliminary data analysis had suggested that these were the 7 αFR mutants with the greatest growth-rate advantage. In strains αFR-2, αFR-8, and αFR-20 we identified mutations in known mating genes: Ste11, Ste5, and Ste7, respectively. In strains αFR-4, αFR-7, and αFR-17 we did not identify any mutations from the yeast tiling arrays; however, subsequent expression analysis suggested that αFR-4 and αFR-7 contain mutations in Ste7 and Ste4, respectively. Sequencing of these genes revealed coding changes in each gene resulting in the following protein modifications: Ste11P656H, Ste7E3ochre, Ste4frameshift, Ste5C198S, and Ste7C198S in αFR-2, -4, -7, -8, and -20, respectively. Fig. 2A shows a mapping of growth-rate coefficients for the identified spontaneous αFR mutants and gene deletions onto the mating pathway; 5 of the 6 faster-growing spontaneous αFR mutants contain a single mutation that reduces basal-signaling-dependent gene expression. Strain αFR-1 is the exception; it is the only instance where we found multiple mutations and mutations outside of the mating pathway. In this strain we found mutations in Apcl (an essential component of the anaphase promoting complex) and Eds1 (an uncharacterized, putative zinc-cluster protein). Apcl and its G1 cofactor Cdh1 play a conserved and critical role in maintaining a G1 arrest, and mutations in Apcl result in premature entry into S phase (15, 16). This strain was selected for αFR resistance, has an expression profile identical to that of canonical ste mutants, and arrests after overexpression of Far1. Because the pheromone signaling pathway is repressed outside G1, mutants that reduce the duration of G1 will reduce the basal expression of mating genes, as APC mutants have been shown to do (17).

We assayed for changes in gene expression for the 7 αFR mutants from the upper end of the growth-rate distribution (s_g > 0), 3 from the lower end of the distribution (s_g = 0), and the targeted gene disruptions (Fig. 2B). The 7 spontaneous αFR mutants from the upper end of the growth-rate distribution can be gained by losing signaling at multiple points in the mating pathway (Fig. S14). All transformed strains arrest after overexpression of the dominant FARI-22 allele (Fig. S1A). The 4 αFR strains not phenotypically characterized are among those that accumulate suppressor mutations and are biased toward the lower end of the growth-rate distribution (Fig. S1B, P = 3.3 x 10^-4, Wilcoxon rank sum).

There are 2 possibilities for the growth-rate advantage observed for αFR mutants: Elimination of basal transcription downstream of Ste12 or elimination of a transient Far1-dependent arrest because of inappropriate activation of the pathway. To distinguish between these possibilities, we measured the growth rate of several targeted gene deletions in the mating pathway. Deletion of the Gβ subunit (Ste4), the MAP kinase kinase (Ste7), or the transcription factor (Ste12) increases growth rate relative to strains in which all 3 genes are intact (Fig. 1B) (P = 2.8 x 10^-5, 5.8 x 10^-3, 2.2 x 10^-6, respectively, t test); however, deletion of the α receptor (Ste2) or Far1 does not (Fig. 1B) [P = 0.23, 0.03, respectively, t test (the value for far1Δ is not significant because we are testing 5 deletion mutants and thus require < 0.05/5 = 0.01 to regard a result as significant)]. This suggests that a growth-rate advantage exists for the subset of sterile strains that abolish basal signaling through the pathway (therefore eliminating basal expression of the mating genes), which depends on Ste4, Ste7, and Ste12, but not Ste2 or Far1 (14).

We phenotypically characterized 41 of the 45 spontaneous αFR mutants that are higher than wild type (P < 0.05, Wilcoxon rank sum test), whereas the distribution of s_g for Canα mutants is similar to wild type (P > 0.05, Wilcoxon). The 8 clones with growth-rate disadvantages of at least 1% were excluded from the statistical analysis. (8) Targeted gene disruptions show that loss of the Gβ subunit (Ste4), the MAP kinase kinase (Ste7), or the transcription factor (Ste12) increases growth rate (P = 2.8 x 10^-5, 5.8 x 10^-3, 2.2 x 10^-6, respectively, t test); however, loss of the receptor (Ste2) or Far1 does not (P = 0.23, 0.03, respectively, t test (the value for far1Δ is not significant because we are testing 5 deletion mutants and thus require < 0.05/5 = 0.01 to regard a result as significant)]. Values represent the mean and standard deviation of 5 to 7 independent gene deletions.

Fig. 1. A subset of αFR mutations provides a competitive growth-rate advantage. (A) Spontaneous αFR mutants have a greater variance of growth-rate coefficients (s_g) and a higher average growth rate than unselected clones (P < 10^-2, Wilcoxon rank sum test), whereas the distribution of s_g for Canα mutants is similar to wild type (P > 0.05, Wilcoxon). The 8 clones with growth-rate disadvantages of at least 1% were excluded from the statistical analysis. (8) Targeted gene disruptions show that loss of the Gβ subunit (Ste4), the MAP kinase kinase (Ste7), or the transcription factor (Ste12) increases growth rate (P = 2.8 x 10^-5, 5.8 x 10^-3, 2.2 x 10^-6, respectively, t test); however, loss of the receptor (Ste2) or Far1 does not (P = 0.23, 0.03, respectively, t test (the value for far1Δ is not significant because we are testing 5 deletion mutants and thus require < 0.05/5 = 0.01 to regard a result as significant)]. Values represent the mean and standard deviation of 5 to 7 independent gene deletions.

the scaffolding protein, Ste5) ultimately leading to a cell-cycle arrest (dependent on Far1) and the induction of mating genes through the transcription factor, Ste12 (13). Expression of mating pathway genes in the absence of pheromone is maintained by basal signaling through the pathway, which depends on the G protein, MAP kinase cascade, and Ste12, but independent of the receptor or Far1.

We phenotypically characterized 41 of the 45 spontaneous αFR mutants by determining the position of each mutation relative to Ste4 and Far1. Overexpression of Ste4 activates the signaling pathway in the absence of pheromone as does expression of a mutant form of Far1 encoded by FARI-22. We transformed each αFR mutant with plasmids containing galactose-inducible STE4 or FARI-22. Five of the 41 successfully transformed strains arrest after Ste4 overexpression indicating a mutation at or before Ste4 in the signal transduction pathway. These 5 mutations are found throughout the distribution of growth-rate coefficients indicating that a growth-rate advantage
The growth-rate advantage of αFα mutants correlates with the elimination of gene expression. (A) A schematic of the pheromone-induced signaling pathway. In a MATA yeast cell, the mating pheromone, αF, binds to the receptor, Ste2. Pheromone-induced signaling activates a heterotrimeric G protein (consisting of Gpa1, Ste18, and Ste4), which in turn, activates a MAP kinase cascade (consisting of the MAP kinase kinase, Ste7, the MAP kinases Fus3 and Kss1, and the scaffolding protein, Ste5) ultimately leading to a cell-cycle arrest dependent on Far1 and a transcriptional response through the transcription factor, Ste12 (13). Expression of mating pathway genes in the absence of pheromone is maintained by basal signaling through the pathway, which is independent of the receptor function (Ste2) or Far1. Identification of the mutations in 5 spontaneous αFα mutants with a growth-rate advantage using yeast tiling arrays (42) identified Ste11P656H, Ste5C198R, and Ste7I406M mutations in strains αF-2, αF-8, and αF-20, respectively. Using this method, we also identified Apc1S838I and Eds1P6L mutations in the strain αF-1. (B) αF strains with a growth-rate advantage reduce gene expression downstream of Ste12, whereas strains without a growth-rate advantage do not show the same reduction in gene expression. Strains are displayed in order of their growth-rate coefficient. The genes displayed are those whose expression changes significantly in the 5 spontaneous αFα mutants with a competitive growth-rate advantage (Fig. S2) or known components of the mating pathway. Some of the apparent down-regulation in strains αF-36 and αF-41 may be an artifact because these strains acquire suppressor mutations that partially restore mating and αF arrest (Fig. S1B). From these arrays, the Ste7I406M and Ste4S838I mutations were identified in strains αF-4 and αF-7, respectively, because the suppression of the expression of specific genes suggests nonsense-mediated decay of the transcripts.

Fig. 2. The growth-rate advantage of αFα mutants correlates with the elimination of gene expression. (A) A schematic of the pheromone-induced signaling pathway. In a MATA yeast cell, the mating pheromone, αF, binds to the receptor, Ste2. Pheromone-induced signaling activates a heterotrimeric G protein (consisting of Gpa1, Ste18, and Ste4), which in turn, activates a MAP kinase cascade (consisting of the MAP kinase kinase, Ste7, the MAP kinases Fus3 and Kss1, and the scaffolding protein, Ste5) ultimately leading to a cell-cycle arrest dependent on Far1 and a transcriptional response through the transcription factor, Ste12 (13). Expression of mating pathway genes in the absence of pheromone is maintained by basal signaling through the pathway, which is independent of the receptor function (Ste2) or Far1. Identification of the mutations in 5 spontaneous αFα mutants with a growth-rate advantage using yeast tiling arrays (42) identified Ste11P656H, Ste5C198R, and Ste7I406M mutations in strains αF-2, αF-8, and αF-20, respectively. Using this method, we also identified Apc1S838I and Eds1P6L mutations in the strain αF-1. (B) αF strains with a growth-rate advantage reduce gene expression downstream of Ste12, whereas strains without a growth-rate advantage do not show the same reduction in gene expression. Strains are displayed in order of their growth-rate coefficient. The genes displayed are those whose expression changes significantly in the 7 spontaneous αFα mutants with a competitive growth-rate advantage (Fig. S2) or known components of the mating pathway. Some of the apparent down-regulation in strains αF-36 and αF-41 may be an artifact because these strains acquire suppressor mutations that partially restore mating and αF arrest (Fig. S1B). From these arrays, the Ste7I406M and Ste4S838I mutations were identified in strains αF-4 and αF-7, respectively, because the suppression of the expression of specific genes suggests nonsense-mediated decay of the transcripts.

significantly decrease the expression of 23 genes but did not significantly increase the expression of any other genes (Fig. S2). The 3 spontaneous αFα mutants from the lower end of the growth-rate distribution and the far1 and ste2 deletions do not decrease the expression of these genes to the degree seen in strains with a growth-rate advantage, consistent with the hypothesis that the growth-rate advantage is the result of elimination of basal expression of the mating pathway genes; the apparent down-regulation of some mating genes observed in strains αFα-36 and αFα-41 could be an artifact because of the acquisition of suppressor mutations that partially restore mating. We conclude that the 2% growth-rate advantage in the most-fit sterile strains is not the result of loss of signaling through the mating pathway per se, but rather the elimination of unnecessary expression from 23 genes related to the mating pathway, supporting the hypothesis that selection, acting on the cost of gene expression, can promote gene loss.

Selection for Mating Efficiency Can Increase the Cost of Gene Expression. Although expressing mating genes decreases growth rate, strains derived from the best characterized laboratory strain, S288c, carry an allele of the Gα subunit (GPA1) that increases basal expression of these genes (18). [The strain used in these experiments, W303, is a mosaic genome comprised of ~85% S288c (Fig. S3A). GPA1 is located on the left arm of chromosome VIII, in a region of predominantly non-S288c descent; however, there is evidence of recombination breakpoints 8 kb upstream and 14 kb downstream of this gene consistent with the idea that the derived allele confers a selective advantage under laboratory conditions (Fig. S3B).] The S288c allele of GPA1 (GPA1Δ-G1406T) contains a T at position 1,406 in place of the ancestral G (found in multiple wild S. cerevisiae isolates and the other sequenced sensu stricto species: S. paradoxus, S. mikatae, S. kudriavzevii, and S. bayanus) resulting in the substitution of isoleucine for serine at amino acid 469, which is located in the C-terminal region of the protein thought to interact with the Ste2 receptor; other mutations in this region constitutively activate the mating pathway (19).

To test the effects of the 2 GPA1 alleles on fitness, we introduced the wild-type GPA1 allele into the laboratory strain and measured the effects of this single mutation on growth rate and mating efficiency relative to the laboratory strain. We isolated 3 individual transformants of each allele (wild-type GPA1 and GPA1-G1406T), which are isogenic except for the GPA1 locus. Expression analysis shows that the wild-type GPA1 allele reduces expression of genes downstream of Ste12 relative to the GPA1-G1406T allele, although these genes are not reduced to the levels seen for sterile mutations (ste7Δ, ste4Δ, and ste12Δ), which eliminate signaling through the mating pathway (Fig. S3). To determine whether reducing the expression of the mating genes by restoring the wild-type GPA1 results in a growth-rate advantage, we measured the relative growth rate by using the FACS-based growth-rate assay and by direct competitions. Strains with the wild-type GPA1 allele have a significant growth-rate advantage over strains with the GPA1-G1406T allele in both assays (Fig. 3B and Fig. S4), sG = 0.92% ± 0.35%, for the FACS-based assay, P = 2.57 × 10^-4, t test; sG = 0.84% ± 0.15% for the direct competition, P < 0.001; the values for sG are in good agreement between these 2 assays, P > 0.10, t test). Consistent with the hypothesis that the growth-rate advantage is the result of reduction in gene expression, we find that, like expression levels, the growth-rate advantage of the wild-type GPA1 allele replacement strains is greater than that observed for strains that do not affect signaling (far1Δ and ste2Δ) and less than that observed for strains that eliminate signaling (ste7Δ, ste4Δ, and ste12Δ) through the mating pathway.

Among the genes up-regulated by the GPA1-G1406T allele are two of the pheromone genes (MFA1 and MFA2); it has been demonstrated that strains producing more pheromone are more attractive to a mating partner (20). To determine whether the 1% growth-rate disadvantage of strains carrying the GPA1-G1406T allele is offset by an increase in mating efficiency, we measured competitive mating efficiency of strains carrying the 2 alleles of GPA1 by using an assay analogous to the competitive growth-rate assay: MATα cells carrying either the wild-type or G1406T allele of GPA1 were mixed and mated to a limited number of
Mating Pathway

Fig. 3. A trade-off between growth rate and mating efficiency. (A) The wild-type allele of GPA1 down-regulates genes in the mating pathway producing an expression profile intermediate to that of deletions eliminating basal signaling (ste1α, ste4α, and ste12Δ) and those not affecting signaling (far1β and ste2Δ). Shown are 3 independent wild-type GPA1 allele replacement strains. (B) Wild-type GPA1 allele replacement strains have a growth-rate advantage relative to the GPA1-G1406T allele strains (ΔGPA1-G1406T = 0.92% ± 0.35% and 0.17% ± 0.34% for the wild-type GPA1 allele and the GPA1-G1406T allele, respectively, p < 2.6 × 10^{-4}, t test). The points represent 3 independent measurements for each of 3 independent transformants of each GPA1 allele. (C) Wild-type GPA1 allele replacement strains have a mating disadvantage relative to the GPA1-G1406T allele strains (Δm = -27.2% ± 6.5%). MATα cells carrying each allele were mixed and allowed to compete for a limiting number of MATα cells. The mating coefficients (Δm) were calculated as the change in the natural logarithm of the ratio of the 2 alleles: Δm = ln(wild-type GPA1/GPA1-G1406T)postmating - ln(wild-type GPA1/GPA1-G1406T)premating.

Discussion

In bacteria, gratuitous gene expression reduces growth rate (21–26). In the most carefully studied case, deregulation of the lac operon can reduce growth rate by ~10% by diverting ribosomes from the synthesis of other proteins (25). This finding implies that any gene slows cell growth in proportion to how strongly it is expressed. We suspect that the cost of gene expression is not specific to bacterial enzymes or genes in the yeast mating pathway, but rather reflects a universal cost of gene expression and that this cost must be borne in all environments where the gene is expressed. In environments where the protein’s expression increases fitness, this cost is offset by larger benefits, but it is never zero. Previous attempts to demonstrate a general fitness advantage from the elimination of dispensable genes in yeast have been unsuccessful (27, 28). Because 97% proteins are expressed at levels <0.1% of total cell protein (29), the growth-rate advantage for the majority of single gene deletions is below the limit of detection by current assays. However, because mating depends on the expression of 23 genes, mutations that eliminate signaling through the mating pathway will eliminate expression of a number of genes and produce a measurable growth-rate advantage.

Given that gene expression is costly, it is surprising that some strains carry a mutation in GPA1, an upstream component of the mating pathway, which increases expression of the mating genes; this polymorphism is one of the strongest trans-acting regulatory polymorphisms between laboratory and wild strains (18). Our results provide a plausible explanation for the existence of the GPA1-G1406T allele: This mutation increases the basal expression of genes in the mating pathway thus increasing mating efficiency by ~30% at a growth-rate cost of ~1% per generation. In the laboratory, cells are mated en masse, a condition that strongly selects for cells that produce more pheromone (20). For these GPA1 variants, where we have determined the effect on growth rate and mating efficiency, we can determine under what conditions each allele will be favored. If these 2 strains are mixed and propagated in a regime where 1 round of mating occurs every 30 generations, these 2 strains would be equally fit (Fig. 4, black trace). If mating is less frequent than every 30 generations, the wild-type allele is favored, whereas if mating is more frequent than every 30 generations, the GPA1-G1406T allele is favored (Fig. 4). During long-term evolution, strains are typically propagated asexually. In this regime, sterile strains, which eliminate basal signaling through the mating pathway, will outcompete mating-proficient strains (Fig. 4, green trace).

The success of different GPA1 alleles in the laboratory versus in wild strains reflects differences in the selective pressures defining fitness in these 2 environments. A similar phenomenon has been observed in E. coli where the laboratory strain B carries an allele of the arginine repressor, argR, that results in deregulation of the arginine biosynthesis pathway, providing a selective advantage under conditions where arginine availability fluctu-
ates rapidly (30). The \textit{argR} \^\textit{B} allele appears to be a naturally occurring variant that is beneficial under laboratory conditions (31). To follow up on the possibility that the GPA1 mutation arose during laboratory cultivation, we traced the lineage of S288c to the wild diploid strain, EM93, which was isolated from a rotting fig near Merced, California in 1938 (32). We sequenced the GPA1 genes in EM93, and found, somewhat to our surprise, that EM93 is homozygous for the GPA1-G1406T allele. Thus, we are left with some ambiguity as to evolutionary origin of this allele: If one discounts a strain mix-up, the mutation appears to have arisen in the wild, raising the question as to how much of laboratory domestication is the result of de novo mutation versus the selection of favorable combinations of naturally occurring alleles.

Gene loss is an important process in evolution. Extensive gene loss in protomitochondria and \textit{Mycobacterium leprae} may have fostered the transition from facultative to obligate intracellular parasites (33), reciprocal gene loss after whole-genome duplication has reinforced species barriers by establishing Dobzhansky-Muller incompatibilities (34), and the loss of key developmental regulators early in vertebrate evolution has been


discussed in ref. 35. Here, we provide evidence for a general cost of gene expression and find that elimination of the expression of 23 genes results in a 2% growth-rate advantage. Assuming that each of these genes contributes equally, the growth-rate advantage attained by eliminating a single dispensable gene is $<0.1\%$. The fate of mutations whose selection coefficient is $>1/N$ is dominated by selection; therefore, for population sizes greater than $\approx 10^3$, such as panmictic microbial populations, selection will oppose unnecessary gene expression, but for small or subdivided populations, drift will dominate for all but the small fraction of strongly expressed genes. Selection for sterile strains during long-term evolution and for the GPA1-G1406T allele supports the hypothesis that selection can optimize the level of gene expression to balance the cost of protein production and the demand for protein function, and argues that proteins that do not increase fitness will be lost.

**Materials and Methods**

**Strains, Plasmids, and Media.** The strains used in this experiment are derived from the base strain, DBY15084, a haploid yeast strain derived from the W303 background with genotype \textit{MATa, ade2-1, can1, his3-11,11, leu2-3,112, trpl-1, ural3}, \textit{bar1::ADE2}, and \textit{hmlA::LEU2}. The full genotypes of all strains used in this study are described in Table S1. Selection for \textit{A\textsuperscript{f}}- and canavanine-resistant clones was performed by spotting plates on media containing 10 \mu g/mL \textit{A\textsuperscript{f}} or 60 \mu g/mL \textit{I}-canavanine as described in ref. 36. Integrative and plasmid transformations were performed by using standard yeast procedures (37). Primers used in this study are described in Table S2. Plasmids pTCN112 and pTCN113 were used for Gal overexpression of \textit{STE4} and the dominant \textit{S2}. Plasmids pTCN112 and pTCN113 were used for Gal overexpression of \textit{STE4} and the dominant \textit{S2}. Plasmids pTCN112 and pTCN113 were used for Gal overexpression of \textit{STE4} and the dominant \textit{S2}. Plasmids pTCN112 and pTCN113 were used for Gal overexpression of \textit{STE4} and the dominant \textit{S2}.

**Competitive Growth-Rate Assays.** Growth rate was determined by using 2 assays: A 

\textit{FACS-based competitive growth-rate assay (39, 40)} and direct competition. For the FACS-based assay, cells were grown to mid-log phase and mixed $\approx 1:1$ with a fluorescently labeled reference strain expressing a Cwp2-YFP fusion protein (40). Cultures were immediately diluted 1:500 into 10 mL of prewarmed YPD (yeast extract, peptone, 2% dextrose) and a 1-ML sample was spun down, resuspended in phosphate buffered saline with 0.1% Tween, and stored at 4 °C. Dilutions were repeated after 12 and 24 h. Samples collected at 0, 12, 24, and 36 h were prepared for FACS by vortex and sonication. The number of YFP-positive (reference) and non-fluorescent (experimental) cells was determined by using an LSR II flow cytometer (BD Biosciences) counting 30,000 total cells for each sample. At each time point a subset of the samples were counted by using a particle counter (Beckman-Coulter) to determine the number of generations between each sample point. The growth-rate coefficient ($s_g$) of each strain relative to the reference was calculated as the rate of the change in the \textit{ln} ratio of experimental to reference versus generations (41). Of the 162 fitness assays performed in this study, only 13 had an $R^2$ value for the correlation between \textit{ln} ratio versus generations of $<0.925$ (Fig. S5). In all cases, 1 of the 4 data points deviated from the trend, most often the result of 1 strain entering a lag phase after mixing. Once the offending data point was removed, $s_g$ and $R^2$ values were recalculated for each experiment. The reference strain has an $\approx 3\%$ disadvantage compared with wild type, therefore the values of $s_g$ were normalized by using the equation $s_g = \frac{\text{Exp} - \text{WT} + \text{Ref} - \text{WT}}{1}$. The unselected clones and the spontaneously occurring \textit{A\textsuperscript{f}}- and canavanine-resistant mutants were normalized to the median $s_g$ of the unselected clones. The targeted gene disruptions were normalized to the median $s_g$ of strains carrying the NatMX marker downstream of the S288c version of GPA1-G1406T.

**In addition to the FACS-based relative growth-rate assay, we used a direct competition to determine the difference in growth rate between strains carrying the GPA1-G1406T and the wild-type allele of GPA1. This assay was performed exactly as the relative growth-rate assay except that the ratio of the 2 strains was determined by quantitative sequencing. Briefly, genomic DNA was prepared from the mixed sample at each time point and the GPA1 allele was amplified and sequenced in both directions. Fluorescence sequencing chromatograms were calculated for the alleles by using peak height as a proxy for abundance. The growth-rate coefficient ($s_g$) of the wild-type GPA1 allele was calculated as the rate of the change in the \textit{ln} ratio of the wild-type GPA1 allele to the GPA1-G1406T allele versus generations.

**Yeast Tilting Arrays and Mutation Identification.** Genomic DNA was prepared from 20 mL of saturated YPD culture by using QIAI\textsuperscript{GEN} Genomic-Tip 100/G and Genomic DNA buffers according to the manufacturer's instructions (QIAGEN). Five micrograms of total genomic DNA were fragmented by sonication (30 0.5-s pulses) by using a microtip sonicator (Misonix) and cleaned by using Zymo Cleanup kit according to the manufacturer's instructions (Zymo Research). Five hundred nanograms of fragmented DNA were labeled by using BioPrime labeling kit according to the manufacturer's instructions with half-volume reactions (Invitrogen). Hybridization to Affymetrix yeast tiling arrays and identification of single nucleotide polymorphisms (SNPs) by using the SNPscanner algorithm were done as described in ref. 42. Seven \textit{A\textsuperscript{f}}- mutants, along with the reference strain DBY15084 were each hybridized to an array. The strain used in this study contains $\approx 8,000$ SNPs relative to the S288c strain on which the array is based (43). The output files from SNPscanner were viewed by using the Integrated Genome Browser software (Affymetrix) and SNPs were calculated for SNPs with the \textit{GPA1} \textit{S2} \^\textit{11} \textit{B} allele that is the SNPs identified from the SNP scanner reference strain. Among the 7 strains tested, 9 SNPs were identified and a 600-bp region centered at each SNP was sequenced from the reference strain and 2 clones from the \textit{A\textsuperscript{f}}- strain in which the SNP was identified. Four of the 9 SNPs were present only in the clone used for array analysis, 5 SNPs were identified in both clones, and none of the SNPs were identified in the reference strain.

**Yeast Gene Expression Microarrays.** Cells were harvested from 10-ML YPD cultures at mid log ($\approx 2 \times 10^7$ cells per mL) by vacuum filtration onto a 25-mm nylon membrane, snap frozen in liquid nitrogen, and stored at $-80^\circ$ C. RNA preparation, labeling, hybridization, and data acquisition were performed as described in ref. 44. Briefly, crude RNA was extracted by the acid phenol lysis protocol and cleaned by using a QIAGEN RNeasy column. Cy3- and Cy5-labeled CTP were incorporated by using the Agilent Low RNA Input Fluorescent Linear Amplification kit. Each Cy5-labeled sample was mixed with the common Cy3-labeled reference strain (DBY15084) and hybridized to an Agilent yeast gene expression 44 k array (Agilent Technologies). Arrays were scanned by using an Agilent DNA microarray scanner and analyzed by using Agilent's Feature Extraction software. Each Agilent yeast gene expression 44 k array contains 7 identical probes for each gene in a given probe set. Combining the final empirical intensities from 7 of the 8 probes for the spontaneous \textit{A\textsuperscript{f}}- mutants gives us 49 estimates for the hybridization intensity in the Cy5 (experimental) and Cy3 (reference) channel. P-values were calculated by using a 2-tailed test of the Cy5 and Cy3 hybridization intensities for each
gene. Significant changes reported in Fig. 3 were determined by eye from a volcano plot of log2(red/green) versus P value and correspond to a 1.5-fold change in expression and P < 10^-3.15 (Fig. S2).

Competitive Mating Assays. Competitive mating assays were performed by competing MATa cells carrying the GPA1-G1406T allele with cells carrying the wild-type allele of GPA1 (marked with NatMX) for a limited number of MATa cells. Each strain was grown to mid-log phase (~10^6 cells per ml) then 5 × 10^6 cells of each MATa strain added to 10 ml of YPD + ADE. A sample was diluted, sonicated, and plated to single colonies on -leu media. MATa mating tester (2 × 10^6 cells) was added to the MATa mixture and filtered onto a 25-mm 0.45-μm nylon filter. Cells were mated on a YPD plate at 30°C. After 5 h, the filters were washed and the cells were diluted, sonicated, and plated to single colonies onto minimal media to select for diploids. Colonies on the -leu and minimal plates were replica plated to YPD plates containing ClonNat to determine the ratio of strains carrying the wild-type and G1406T alleles before and after mating. The mating coefficient (s_m) was calculated as the change in the natural logarithm of the ratio of the 2 alleles: s_m = ln(wild-type GPA1/ GPA1-G1406T)postmating − ln(wild-type GPA1/GPA1-G1406T)premating.

Notebook. The complete laboratory notebook describing these experiments is available at http://www.genomics.princeton.edu/glang/notebooks.htm.

Acknowledgments. We thank M. Desai, L. Kruglyak, and members of the Botstein and Murray labs for comments on the manuscript; J. Shapiro and L. Kruglyak (Princeton University, Princeton, NJ) for providing the wild-type allele; B. Tilton and C. DeCoste for assistance with flow cytometry; and A. Ward, D. Gresham, J. Buckles, and D. Storton for assistance with microarrays. This work was supported by National Institute of General Medical Sciences Centers of Excellence Grants GM068763 (to A.W.M.) and GM071508 (to D.B.) and National Institutes of Health Grants GM43987 (to A.W.M.) and GM046406 (to D.B.).