Review

Experimental genomics of fitness in yeast
Graham Bell*

Biology Department, McGill University, 1205 Avenue Docteur Penfield, Montreal, Quebec H3A 1B21, Canada

The set of single-gene deletions in yeast can be used to evaluate the effect of mutation on fitness over the whole genome. The measurement of growth in pure culture or relative growth in mixtures has confirmed that most deletions have little effect in laboratory culture. Moreover, there is a sharp distinction between lethality and a very mild impairment of growth, with very few intermediate cases. Different components of fitness, such as growth rate and yield, are positively correlated. Growth is also positively correlated across environments, although new conditions of growth usually identify a few conditionally impaired strains. Double mutants on average show alleviating epistasis, although a few per cent of combinations are synthetic lethal. The properties of the yeast deletion set provide us with the first genome-wide account of fitness, although transferring these conclusions to the field is a task for the future.

Keywords: gene deletion; fitness components; epistasis; environmental heterogeneity; plasticity

1. THE UNEXPECTED MILDNESS OF MUTATION

The classical view of mutations, that held sway throughout the twentieth century, is that almost all mutations are deleterious, in proportion to the magnitude of genetic disruption. The geometrical analogy developed by Fisher (1930) was particularly influential. It depicts the optimal state of the population as a point in space representing a combination of character values. The current state of the population is a second point, more or less distant from the optimum. Any random change in a character, brought about by mutation, will move the population either closer to the optimum or farther away. Large random changes are almost certain to relocate the population farther away, and will be prevented from spreading by strong purifying selection. Only small changes have any appreciable chance of being beneficial, and adaptation will therefore involve the successive fixation of a long series of mutations of small effect.

In the last few years it has become possible to estimate the effects of mutation on fitness in a systematic fashion, throughout the entire genome. The results have been unexpected. Early studies using gene disruption by transposition revealed mostly mild effects, with fitness being reduced by only a few per cent. When it became possible to delete large numbers of genes by homologous recombination in yeast, it was immediately apparent that most deletions have little effect in laboratory culture. Moreover, there is a sharp distinction between lethality and a very mild impairment of growth, with very few intermediate cases. Different components of fitness, such as growth rate and yield, are positively correlated. Growth is also positively correlated across environments, although new conditions of growth usually identify a few conditionally impaired strains. Double mutants on average show alleviating epistasis, although a few per cent of combinations are synthetic lethal. The properties of the yeast deletion set provide us with the first genome-wide account of fitness, although transferring these conclusions to the field is a task for the future.

Keywords: gene deletion; fitness components; epistasis; environmental heterogeneity; plasticity

2. MATERIAL AND METHODS

(a) Strains

The haploid MATa yeast single deletion set was obtained from Open Biosystems (www.openbiosystems.com). The ancestor is BY4741 (MATa his3Δ1 leu2Δ3 met15Δ0 ura3Δ0).

(b) Media

YPD is yeast peptone dextrose medium: 20 g peptone, 10 g yeast extract, 20 g dextrose to 1 l with distilled water. Minimal medium: 1.4 g yeast nitrogen base (without amino acids), 20 g glucose to 1 l with distilled water. Low-N minimal medium has 1.4 g yeast nitrogen base. Maple and birch sap was obtained by coring trees at Mont St-Hilaire, Quebec, in March 2005, sterile-filtering the homogenized samples and storing frozen. Bark and leaf infusions were made by steeping a given quantity of plant material in an equal volume of water at room temperature for 24 h and sterile-filtering the supernatant. All cultures were incubated at 28°C except in the high temperature experiment where they were incubated at 37°C.

Received 18 November 2009
Accepted 12 January 2010

This journal is © 2010 The Royal Society
(c) Pure-culture assays
For pure-culture assays, all strains were replated to 96-well plates each holding 48 deletion strains and 12 copies of the ancestor in interior wells, with all marginal rows blank except for plate identification inoculations. For each cycle of growth, wells were inoculated with 10 μl of stock culture and cultured for 48 h at 28°C, before a second set was inoculated from the first and again grown for 48 h at 28°C before being stirred and optical density recorded (FluoStar Optima, BMG Lab Technologies). Transfers were done in sterile conditions using a Biomek FX automated liquid handling system inside a Baker BioProtect II laminar flow hood. Plates were routed with a robotic handler (Beckman Coulter).

3. HOW DOES SELECTION PRESERVE GENETIC INTEGRITY?
The problem to be solved is: how can the mild phenotype of most deletion strains be reconciled with the maintenance of fully functional gene sequences? There are three leading explanations.

— **Chronic weak purifying selection.** Loss-of-function mutations are not neutral, but rather mildly deleterious, and are therefore prevented from spreading by purifying selection. They will continue to be maintained at low frequency by mutation pressure, but any given individual will bear intact sequences throughout almost its whole genome. If this explanation is correct, then careful studies will show that the growth and fitness of viable deletion strains always falls short of their intact ancestor.

— **Frequent strong synthetic selection.** The genome is moderately redundant, in the sense that the failure of one activity can be compensated by the extension of another. Hence, the loss of a single gene can be tolerated because another gene can take its place. If this second gene should also be lost, however, fitness will be severely reduced. Hence, genetic integrity is preserved because a lineage bearing a loss-of-function mutation at one locus will sooner or later be doomed by a loss-of-function mutation at a second locus. This explanation requires that neutral single-locus deletions should often combine to give a synthetic lethal effect in the double deletant.

— **Intermittent strong directional selection.** This mildness of most deletions is attributable to highly permissive conditions of growth. In more stressful conditions, more loci would be lethal when deleted. Moreover, a different set of loci is essential for each distinctive set of conditions. Genetic integrity is preserved in a changing environment because any lineage bearing a loss-of-function mutation at any locus will be eliminated when the conditions for which that locus is essential appear. This explanation requires that single-locus deletions that are neutral in some environments should be severely deleterious in others.

I shall now summarize the evidence for each of these three processes. They are not mutually exclusive, and each may be responsible, to some degree, for the maintenance of genetic integrity and normal function. I shall then use genomic data to evaluate the contribution of each to selection in natural populations.

4. CHRONIC WEAK PURIFYING SELECTION

(a) **Yield in pure culture**
The most straightforward way of estimating the average effect of gene deletion is to compare the growth of each strain in pure culture with the ancestral value. I grew all 4846 viable haploid strains of the yeast single deletion set in such a way that each strain could be compared with two replicate cultures of the ancestral undeleted strain growing on the same plate (see §2). The measure of growth was optical density of the stirred culture after 48 h of growth at 28°C in rich medium (YPD). Standard yield was then calculated as $Y_{\text{standard}} = \frac{\text{strain} - \text{ancestors}}{\text{blank}}$, using the average of uninoculated wells on each plate as a blank. A score of 0 indicates that the deletion has no effect, whereas a score of −1 indicates that it is lethal. The expected outcome on the Fisherian view is an exponential decline in frequency from a mode of severely impaired strains with large negative $Y_{\text{standard}}$. Instead, $Y_{\text{standard}}$ was narrowly and almost symmetrically distributed around a mode near zero (figure 1). A few strains (about 20) were nearly lethal and grew only very slowly; only a very few strains (about 10) had intermediate values. Excluding these cases, the phenotypic standard deviation of the means was $\sigma_p = 0.052$ and the genetic standard deviation $\sigma_g = 0.039$. This shows that almost all genes in the yeast genome genes fall into two sharply delineated categories, defined by the effect of deletion: Fisherian genes, whose deletion is lethal, and non-Fisherian genes, whose deletion is more or less inconsequential. Moreover, the overall effect of deletion was beneficial: $Y_{\text{standard}}$ exceeded zero for about two-thirds of all loci (3489/4846), with a mean value of +0.0147 (s.d. = 0.084, s.e. = 0.0012; mean excluding the few nearly lethal strains is +0.0195, s.d. = 0.052, s.e. = 0.0007).

To construct the deletion set, the resident gene is displaced by a construct including a *kanMX* deletion cassette conferring resistance to geneticin. The effect on mean yield might therefore be attributable either to the removal of a given gene, or to the insertion of the construct. I identified 30 open reading frames (ORFs) in the *Saccharomyces* genome database (www.yeastgenome.org), which are annotated as ‘Dubious ORF unlikely to encode a protein, based on available experimental and comparative sequence data’, and which do not overlap any other ORF. Deleting these ORFs should have no effect on fitness. The mean standard yield of strains bearing deletions of dubious ORFs in YPD was +0.0350 (s.e. = 0.0074), compared with +0.0147 (s.e. = 0.0012) for all ORFs. I also identified nine genes associated with the sexual cycle, which are unlikely to be expressed in vegetative growth and whose deletion should likewise be neutral. The mean standard yield of these strains in YPD was +0.0329 (s.e. = 0.0085). Hence, the elevation of growth caused by deleting a non-coding or non-expressed ORF is similar to, or greater than, the effect of deleting a coding gene. The general elevation of fitness in YPD associated with gene deletion is therefore likely to be attributable to the insertion of the cassette. If the effects of gene deletion and cassette insertion are additive, then the average effect of gene deletion on growth in YPD is $0.0147 - 0.0350 = -0.0203$. Hence, gene deletion causes a marginal loss of yield of about 2 per cent on average.
The mildness of mutation in rich medium may be misleading if deletion causes a much greater loss of fitness in more stressful conditions, such as are presumably typical of cultures no change in the frequency of wild-type was observed, whereas in 133 cultures the deletants increased. Only 74 different strains were represented among the successful deletants, and some appeared to have hitch-hiked with beneficial mutations arising during the course of the competition; consequently, only 12 consistently superior deletants could be identified.

Breslow et al. (2008) monitored the outcome of competition between intact and deleted strains by fluorescent-activated cell sorter, using contrasted fluorescent protein markers. They reported that about 45 per cent of deletions were associated with impaired fitness in minimal medium. A few deletion strains were consistently superior, many of which involved genes in the nutrient-sensing protein kinase pathway. Reducing the level of expression of essential genes by disrupting the 3′ untranslated region with a marker cassette also caused reduced fitness for about 40 per cent of loci.

**Correlation among fitness components in rich medium**

Yield, growth rate and competitive ability may all contribute to fitness, in proportions that will depend on conditions of culture. They are all positively correlated, to a degree limited by repeatability. Yield and growth rate have $r = +0.44$. Both are correlated with competitive ability. For the deletion-pool studies by Giaever et al. (2002); Steinmetz et al. (2002) and Deutschbauer et al. (2005), the correlations of standard yield with selection coefficient are $+0.13$, $+0.37$ and $+0.37$, respectively; for growth rate, the corresponding values are $+0.13$, $+0.66$ and $+0.56$.

Hence, these large-scale studies lead to the simple conclusion that the effect of a loss-of-function mutation at a locus is similar for all the main components of fitness, at least for cultures grown in rich medium in the laboratory.

**Relative fitness in stressful media**

The mildness of mutation in rich medium may be misleading if deletion causes a much greater loss of fitness in more stressful conditions, such as are presumably typical of cultures no change in the frequency of wild-type was observed, whereas in 133 cultures the deletants increased. Only 74 different strains were represented among the successful deletants, and some appeared to have hitch-hiked with beneficial mutations arising during the course of the competition; consequently, only 12 consistently superior deletants could be identified.

Breslow et al. (2008) monitored the outcome of competition between intact and deleted strains by fluorescent-activated cell sorter, using contrasted fluorescent protein markers. They reported that about 45 per cent of deletions were associated with impaired fitness in minimal medium. A few deletion strains were consistently superior, many of which involved genes in the nutrient-sensing protein kinase pathway. Reducing the level of expression of essential genes by disrupting the 3′ untranslated region with a marker cassette also caused reduced fitness for about 40 per cent of loci.

**Correlation among fitness components in rich medium**

Yield, growth rate and competitive ability may all contribute to fitness, in proportions that will depend on conditions of culture. They are all positively correlated, to a degree limited by repeatability. Yield and growth rate have $r = +0.44$. Both are correlated with competitive ability. For the deletion-pool studies by Giaever et al. (2002); Steinmetz et al. (2002) and Deutschbauer et al. (2005), the correlations of standard yield with selection coefficient are $+0.13$, $+0.37$ and $+0.37$, respectively; for growth rate, the corresponding values are $+0.13$, $+0.66$ and $+0.56$.

Hence, these large-scale studies lead to the simple conclusion that the effect of a loss-of-function mutation at a locus is similar for all the main components of fitness, at least for cultures grown in rich medium in the laboratory.

**Relative fitness in stressful media**

The mildness of mutation in rich medium may be misleading if deletion causes a much greater loss of fitness in more stressful conditions, such as are presumably typical of cultures no change in the frequency of wild-type was observed, whereas in 133 cultures the deletants increased. Only 74 different strains were represented among the successful deletants, and some appeared to have hitch-hiked with beneficial mutations arising during the course of the competition; consequently, only 12 consistently superior deletants could be identified.

Breslow et al. (2008) monitored the outcome of competition between intact and deleted strains by fluorescent-activated cell sorter, using contrasted fluorescent protein markers. They reported that about 45 per cent of deletions were associated with impaired fitness in minimal medium. A few deletion strains were consistently superior, many of which involved genes in the nutrient-sensing protein kinase pathway. Reducing the level of expression of essential genes by disrupting the 3′ untranslated region with a marker cassette also caused reduced fitness for about 40 per cent of loci.

**Correlation among fitness components in rich medium**

Yield, growth rate and competitive ability may all contribute to fitness, in proportions that will depend on conditions of culture. They are all positively correlated, to a degree limited by repeatability. Yield and growth rate have $r = +0.44$. Both are correlated with competitive ability. For the deletion-pool studies by Giaever et al. (2002); Steinmetz et al. (2002) and Deutschbauer et al. (2005), the correlations of standard yield with selection coefficient are $+0.13$, $+0.37$ and $+0.37$, respectively; for growth rate, the corresponding values are $+0.13$, $+0.66$ and $+0.56$.

Hence, these large-scale studies lead to the simple conclusion that the effect of a loss-of-function mutation at a locus is similar for all the main components of fitness, at least for cultures grown in rich medium in the laboratory.

**Relative fitness in stressful media**

The mildness of mutation in rich medium may be misleading if deletion causes a much greater loss of fitness in more stressful conditions, such as are presumably typical of cultures no change in the frequency of wild-type was observed, whereas in 133 cultures the deletants increased. Only 74 different strains were represented among the successful deletants, and some appeared to have hitch-hiked with beneficial mutations arising during the course of the competition; consequently, only 12 consistently superior deletants could be identified.
natural environments. I tested the haploid deletion set in eight more stressful media: minimal, low-N minimal, YP ethanol, high temperature, tree sap and infusions of leaves and bark (see §2). Average yield relative to the intact ancestor was greater in some conditions and less in others. The deletion strains were most severely impaired, however, in the most stressful environments, leaf and bark infusions, which had \( Y_{\text{standard}} = -0.55 \) and \(-0.20\), respectively. Jasnos et al. (2008) found that the effect of deletion on growth rate was actually alleviated in more stressful environments (minimal, high temperature, caffeine and salt). Growth rate and yield were affected in the same direction by culture in minimal medium and at high temperature. Hence, there is no simple or consistent tendency for stress to exacerbate the effect of gene deletion on fitness.

5. FREQUENT STRONG SYNTHETIC SELECTION

All the experiments reviewed above tested the deletion set in haploid or homozygous state. In practice, interactions between allelic and non-allelic genes are likely to be pervasive, and both dominance and epistasis will contribute to the effectiveness of selection.

(a) Dominance

Purifying selection will be most effective for mutations that are dominant for fitness. Steinmetz et al. (2002) provide competition pool data for both homozygous and heterozygous deletion strains. Some heterozygous strains are impaired, but most have nearly average fitness, even for deletions that substantially reduce competitive growth as homozygotes. Hence, most deletions are almost completely recessive (figure 4). Phadnis & Fry (2005) claim that there is a strong and consistent relation between dominance and competitive growth, but this is attributable to autocorrelation between their measures of dominance and selection coefficients. Being recessive, most deletions will be sheltered from selection in wild diploids with infrequent sexual episodes.

(b) Epistasis

Mutations will be removed more effectively if they are more deleterious in combination than when they are single. In the extreme case, loss-of-function mutations at two different loci might be neutral when either is borne, yet lethal when both are borne together. Synthetic lethal interactions will prevent loss-of-function mutations from accumulating. The balance between single-locus and synthetic effects depends on the genomic mutation rate. In small genomes, with low overall mutation rates, most genetic deaths are caused by the single-locus effect. In larger genomes, synthetic effects contribute an increasing fraction of genetic deaths, and become the predominant source of purifying selection when the genomic mutation rate exceeds about one per replication.

Tong et al. (2004) reported a systematic survey of 132 focal single-deletion strains, each of which was crossed with about 5000 other strains in order to isolate the double deletion strains. This generated about \(10^5\) double deletion strains, of which about 4000 had synthetic lethal or synthetic sick phenotypes. Hence, synthetic lethality will give rise to selection of a few percent against deletions that are individually neutral.

More generally, genetic interactions may be alleviating (positive epistasis) or aggravating (negative epistasis). With alleviating epistasis, the fitness of a double mutant is less than the product of the fitnesses of the single mutants; with aggravating epistasis, the fitness of a double mutant is less than the product of the fitnesses of the single mutations; thus, synthetic lethality is the extreme case of aggravating epistasis. Jasnos & Korona (2007) dissected tetrads from 639 crosses between single-deletion strains with low maximum growth rate to isolate the nil-mutation (00), single-mutation (01,10) and double-mutation (11) segregants. Epistasis for exponential growth rate \( m \), estimated as \( e = (m_{00} + m_{11}) - (m_{01} + m_{10}) \), was distributed roughly around zero, but with an excess of positive values at low overall growth rates and a slight excess of negative values at high growth rates (figure 5). The excess of alleviating epistasis is maintained in stressful conditions (Jasnos et al. 2008). It should be noted, however, that systematic synthetic screens (such as that reported by Tong et al. 2004) suggest an excess of aggravating epistasis, primarily among strains with high growth rates. I conclude that genetic interactions are pervasive, but are probably not predominantly aggravating and so do not on balance facilitate the elimination of deleterious mutations.
6. INTERMITTENT STRONG DIRECTIONAL SELECTION

(a) Genetic correlation across environments

The deletion strains do not respond uniformly when growth conditions change. I estimated yield in eight culture media. Four were simple laboratory media: YPD, glucose minimal medium, ethanol-supplemented medium and YPD at 37°C. The other four were natural fluids or extracts: the sap of maple and birch, and infusions of oak bark and leaves. Estimates of variance components showed that genotype-environment interaction \( \sigma^2_{GE} \) was comparable in magnitude to the genetic variance \( \sigma^2_G \) itself: \( \sigma^2_G/(\sigma^2_G + \sigma^2_{GE}) = 0.39 \). The correlation over environments, however, was positive in most (92/112) pairwise comparisons of replicate trials in different media, with an average of \( r = +0.042 \) (s.e. = 0.005). As the correlation of replicate trials in the same medium was only \( r = +0.185 \), this suggests a true correlation over environments of about +0.2. Giaever et al. (2002) conducted trials in six growth conditions, with higher repeatability, leading to a comparable value of average correlation of \( r = +0.269 \) (s.e. = 0.045).

Other authors have tested a much wider range of simpler manipulations by adding cytotoxic agents to standard growth media. Brown et al. (2006) tested 4756 genes in 51 media and found average \( r = +0.0979 \). Parsons et al. (2006) tested 3418 genes in 82 media and obtained a very similar estimate of average \( r = +0.0997 \). The distribution of the cross-environment genetic correlation is strikingly similar in these studies: 75–80% of estimates positive, modal between 0 and +0.1, and with a long right-hand tail of positive values (figure 6).

(b) Mean fitness and plasticity

The variance of growth across environments reflects the ability of a strain to maintain a consistent level of fitness. Strains with lower variance are more plastic, in that they are able to express a similar phenotype in different conditions. The relationship between the mean and the variance of growth across environments will influence overall fitness. Since the most fit combination would be high average growth and a low variance of growth, it has been suggested that there should be a trade-off between mean fitness and plasticity, leading to a positive relationship between the mean and the variance of growth. The actual relationship is a striking ‘shellburst’ plot that roughly fills in a parabolic region of parameter combinations (figure 7). The growth of mediocre strains in different media is distributed more or less symmetrically around the population mean, but extreme strains grow exceptionally well (or poorly) in several media, reflecting the positive correlation across media. This produces a distribution which is still modal around the population mean, but has a long left or right tail, and consequently has greater variance. One consequence of this pattern is that there are no strains that have both high average fitness and high plasticity.

(c) Conditional lethals

Although the growth of most viable deletion strains is weakly positively correlated across environments, this does not exclude the possibility that any given strain may be lethal in a small fraction of environments. In any given environment, then, a few such conditional lethals would be detected, and if enough environments were screened, the great majority of deletions would be found to be conditionally lethal.

For the eight media I tested, I found that the average number of new conditional lethals identified was about 25 per new medium. In the more extensive surveys reported by Brown et al. (2006) and Parsons et al. (2006), the response variable is the change in frequency in pool competition trials, and there is no discrete class of lethals. I have arbitrarily defined subvitals as those strains falling 2 s.d. units or more below the mean fitness over all strains. If there is no genetic correlation between media, the overall frequency of conditional subvitals in a random set of \( E \) media is \( \Sigma p(1 - p) \), where \( p \) is the frequency of subvitals in a single medium, and the summation is taken from \( i = 1 \) to \( i = E - 1 \). Plots of the cumulative frequency of conditional sublethals for an increasing number of media fall below expectation because of the positive genetic correlation between media, and approach a value of about 0.5 for 50–80 media (figure 8). Hence, about half of all deletion strains will be severely deleterious when exposed to one or more agents. The average number of new conditional sublethals uncovered by each new medium in these two surveys was 37 (Brown et al. 2006) and 22 (Parsons et al. 2006).

The most extensive survey reported to date was by Hillenmeyer et al. (2008), who screened all viable deletion strains in media each containing one of 178 different small molecules. Among 3853 homozygous deletion strains with normal growth in complete medium (according to the Deutschbauer dataset), 3437 are substantially impaired in at least one of the experimental media. Most of the loci with no phenotype in any medium do not encode a protein, or are only weakly expressed, if at all. Hence, this study claims to demonstrate that every gene is necessary for normal growth in some conditions. It should be noted, however, that many of the molecules used are biologically active and may inhibit particular
proteins, so the phenotypes that are expressed may be closer to synthetic genetic interactions than to conventional genotype–environment interaction. The number of conditional lethals exposed on average by each new molecule was \(3437/178\)\(^{25}\)\(^{19}\).

In short, these surveys consistently report that altering a single factor in the physical conditions of growth uncovers about 20 additional loci, whose deletion is lethal or severely deleterious. This consistency is all the more remarkable, perhaps, in view of the lack of any objective definition of what constitutes an environmental factor.

7. DISCUSSION: HOW DOES SELECTION MAINTAIN GENETIC INTEGRITY IN NATURE?

Studies of the yeast deletion set have established or confirmed a series of broad generalizations about the variation of fitness.

- Most mutations have only a slight effect on fitness, even when they completely disrupt a gene.
- The effect of mutation is not in general exacerbated by stressful conditions.
- Loss-of-function mutations are recessive.
- A few per cent of pairwise combinations of viable mutations are synthetically lethal.
- Among viable combinations, both aggravating and alleviating interactions are found, with an excess of alleviating interactions.
- The correlation of fitness across environments is positive.
- Plasticity is related to average fitness by a ‘shellburst’ plot in which strains with high average fitness and high plasticity are rare or absent.
- When exposed to any novel environmental factor, a small number of loci (perhaps about 20 on average) are conditionally lethal.

These are general features of variation over the whole genome, insofar as variation can be evaluated through the properties of deletion strains. They are not necessarily features of variation in evolved populations. To show how arbitrary and evolved populations may differ, consider the growth of the strains in two different media. This will usually give a scatter plot with a slight positive correlation. We might argue, however, that only strains that are able to grow fairly well in both media are of interest, so the analysis should be restricted to the few percent of strains with the highest mean fitness across media. The plot now has a strong negative correlation. In one sense, this is an artefact: plotting the best 1 per cent of two large series of independent normal random numbers will yield a regression slope of \(-1\) (if the two series have equal variance) and \(r^2 \approx 0.5\). In another sense, it is what is to be expected as the consequence of selection, through the elimination of strains inferior in both environments. The statistical features of the deletion strains provide a reliable basis for characterizing genetic variation, but will often be modified by selection either in natural or in experimental populations.

(a) Fitness of deletion strains in nature

The effect of gene deletion on fitness in natural conditions of growth has not yet been evaluated by field
experiments. Broad surveys of genetic variation in wild yeast (Liti et al. 2009; Schacherer et al. 2009) have identified strains with large deletions which probably have a very similar effect on fitness as the precise and complete deletions of the laboratory strains. Their occurrence proves that some deletions, at least, have only mild effects on fitness in the wild. Schacherer et al. list 254 large deletions collected from 63 Saccharomyces cerevisiae strains. When cultured in YPD, the standard yield of the deletion strains corresponding to these isolates is significantly higher in my experiments than the average over all viable deletion strains (0.0298 versus 0.0147; \( t = 4.40, \text{d.f.} = 5009, p < 0.001 \)). The pool competition experiments of Giaever and Deutschbauer give a similar result. The mean score of the isolates over all media in the experiment by Parsons et al. (2006) is higher than the overall mean of all viable deletion strains (0.0816 versus 0.0048; \( t = 6.57, \text{d.f.} = 4250, p < 0.001 \)), and the variance is less (0.1437 versus 0.2018; \( t = -3.47, \text{d.f.} = 4250, p < 0.001 \)). From the experiment by Brown et al. (2006), the isolates do not exceed the mean score of all strains (−0.039 versus −0.026; \( t = -2.28, \text{d.f.} = 4435, 0.05 > p > 0.02 \)), although the variance is again less (0.041 versus 0.071; \( t = -4.44, \text{d.f.} = 4435, p < 0.001 \)). Hence, the loci found to be partly or wholly deleted in wild isolates tend to have greater average fitness and a lower variance of fitness over media when deleted and tested in the laboratory.

The frequency of isolates in which a given locus is deleted might provide a rough quantitative measure of its fitness in nature. This frequency is uncorrelated with growth in YPD in any of the data sets. It is weakly correlated (\( r = 0.205, \text{d.f.} = 138, p = 0.01 \)) with the average score over all media in Parson’s experiments, and very weakly correlated with the variance of growth over media (\( r = -0.025, \text{n.s.} \)). The appropriate fitness measure for a lineage experiencing a succession of different conditions is the geometric mean, which is approximately equal to the arithmetic mean minus half the environmental variance. This is positively correlated over loci with the frequency of sites in which the locus is deleted (\( r = +0.26, p < 0.01 \); figure 9). The relationship appears to be triangular rather than linear, in that genotypes with low geometric mean fitness are seldom found. Brown’s data yields a similar triangular plot, but with no linear regression.

Most deletions found in several or many isolates are spurious ORFs, genes of unknown function or members of gene families with several similar copies in the genome (figure 9). This is easily explained if the deletions found commonly in natural populations are simply those for which purifying selection is least effective. Nevertheless, deletions involving 120 annotated genes were identified in 1–15 isolates, comprising 18 per cent of all the deletions found. They had a mean score in Parsons’ survey scarcely inferior to that of spurious or unannotated ORFs (0.067, s.d. = 0.139). Hence, genomes in natural populations can sustain loads of up to a dozen or so deletions of functional genes.

(b) Fitness in complex and variable environments

Yeast is normally cultured in rich, undefined media such as YPD, or in minimal medium with a single limiting carbon source such as glucose. Natural environments are unlikely to be similar to either of these conditions, because they will vary in time and space. It is conceivable that there is a single limiting factor at any given time whose nature frequently changes. This would eliminate loss-of-function mutations at about 20 loci at a time. This seems very unlikely, since natural isolates have not accumulated loss-of-function mutations for a large fraction of the genome.

It is more likely that natural environments contain a variety of substrates at low concentrations, none able alone to sustain growth; they will also contain a range of toxins, growth inhibitors and waste products emitted by other organisms, primarily microbes, also individually at

![Figure 8. The cumulative frequency of conditional sublethal loci when the deletion set is tested over a large number of substrates.](https://example.com/figure8)

![Figure 9. Frequency of deletions among natural isolates in relation to geometric mean fitness in culture.](https://example.com/figure9)
low concentration. The particular range of substrates and toxins may change continuously over time, as the result of seasonal and year-to-year changes in physical processes and community composition. The meagre experimental evidence that is available suggests that selection in complex environments leads to the evolution of overlapping incomplete generalists, each of which can exploit or tolerate a substantial fraction of the prevailing conditions (Barrett et al. 2005). Loss-of-function mutations that narrow the niche of a lineage would be deleterious because each would reduce its rate of growth or limiting abundance by an amount that would be inversely proportional to niche width. This pattern might tend to evolve if the cross-environmental correlation of mutations is positive, such that some combinations of mutations are impaired in given conditions, whereas others are unaffected. It would also be facilitated by the shellburst relationship between plasticity and mean fitness, with mutations having little effect in most conditions but having much higher or much lower fitness in others.

In terms of the three options identified in §1, the most plausible explanation for the preservation of genetic integrity over evolutionary time seems to be chronic weak purifying selection elicited by the complexity of natural environments. Strong genetic interactions and frequent changes in conditions seem less likely to be important.

This scheme seems consistent with laboratory observations and with the occurrence of modest frequencies of deletions of annotated genes in natural populations. However, it is far from being securely established. First, there is no entity that unequivocally constitutes a unit environmental factor, in the same way that a gene constitutes a unit genetic factor. Small molecules are an interesting possibility, but their effect will be modulated by the physical context, such as pH and moisture, and they do not fully represent vicissitudes such as being eaten by a nematode. Secondly, we know very little about the natural conditions of growth of wild yeast. It can be reliably isolated from plant surfaces such as tree bark and leaves (Naumov et al. 1998; Glushakova et al. 2007; Replansky 2007), but its physical and chemical conditions of life, and its interactions with the community in which it is embedded, are very poorly understood. Finally, field experiments have not yet been attempted because suitable experimental material is not available. The deletion strains cannot be deployed in the field because they are constructed in a multiply auxotrophic background, and no comparable resource is yet available for wild yeast. The whole-genome surveys that yeast bioinformatics has made possible have indeed provided us with unparalleled insight into the variation of fitness. The interpretation of these surveys in relation to the conditions in which yeast has evolved awaits the development of an ecological and evolutionary yeast model that will be as powerful as its laboratory precursor (Replansky et al. 2008).

This research was funded by the Natural Sciences and Engineering Research Council of Canada and the Canadian Foundation for Innovation.

REFERENCES


Fudala, A. & Korona, R. 2009 Low frequency of mutations with strongly deleterious but non-lethal fitness effects. Evolution 53, 2164–2171.


Replansky, T. 2007 Saccharomyces sensu stricto as a model system in ecology and evolution. MSc thesis, McGill University, Montreal, Canada.

Replansky, T., Koufopanou, V., Greig, D. & Bell, G. 2008 Saccharomyces cerevisiae as a model system in ecology

This article was first presented as the Presidential Address to the annual meeting of the Canadian Society for Ecology and Evolution in Vancouver, May 2008. The robot experiments were run by Zhang Ming Wang. I am grateful to Austin Burt (Imperial College, London) and Corey Nislow (University of Toronto) for comments on the manuscript.


