

Synonymous mutations in representative yeast genes are mostly strongly non-neutral

<https://doi.org/10.1038/s41586-022-04823-w>

Xukang Shen¹, Siliang Song¹, Chuan Li^{2,3} & Jianzhi Zhang^{1✉}

Received: 3 August 2021

Accepted: 28 April 2022

Published online: 8 June 2022

 Check for updates

Synonymous mutations in protein-coding genes do not alter protein sequences and are thus generally presumed to be neutral or nearly neutral^{1–5}. Here, to experimentally verify this presumption, we constructed 8,341 yeast mutants each carrying a synonymous, nonsynonymous or nonsense mutation in one of 21 endogenous genes with diverse functions and expression levels and measured their fitness relative to the wild type in a rich medium. Three-quarters of synonymous mutations resulted in a significant reduction in fitness, and the distribution of fitness effects was overall similar—albeit nonidentical—between synonymous and nonsynonymous mutations. Both synonymous and nonsynonymous mutations frequently disturbed the level of mRNA expression of the mutated gene, and the extent of the disturbance partially predicted the fitness effect. Investigations in additional environments revealed greater across-environment fitness variations for nonsynonymous mutants than for synonymous mutants despite their similar fitness distributions in each environment, suggesting that a smaller proportion of nonsynonymous mutants than synonymous mutants are always non-deleterious in a changing environment to permit fixation, potentially explaining the common observation of substantially lower nonsynonymous than synonymous substitution rates. The strong non-neutrality of most synonymous mutations, if it holds true for other genes and in other organisms, would require re-examination of numerous biological conclusions about mutation, selection, effective population size, divergence time and disease mechanisms that rely on the assumption that synonymous mutations are neutral.

The cracking of the genetic code in the 1960s revealed that between one-quarter and one-third of single nucleotide mutations in protein-coding genes do not alter protein sequences^{1,2}. Although these synonymous mutations are not strictly neutral, because they could influence many processes^{6–8} such as transcription factor binding⁹, transcription¹⁰, pre-mRNA splicing⁷, mRNA folding¹¹ and stability^{12,13}, translational initiation¹⁴, efficiency^{15,16} and accuracy^{17,18}, and co-translational protein folding^{19,20}, the vast majority of synonymous mutations are presumed to be at least nearly neutral^{1–5}, in contrast to nonsynonymous mutations, which alter protein sequences and frequently affect fitness^{3–5}. The neutrality or near neutrality of synonymous mutations is widely assumed in inferring mutation rate, pattern and mechanism, testing natural selection, estimating effective population sizes (N_e) and neutral genetic diversities commonly considered in conservation policy making in addition to population and evolutionary biology, and dating evolutionary events such as population or species divergences and gene or genome duplication^{3–5}. This assumption also diverts the mechanistic study of disease from synonymous mutations²¹.

Nevertheless, some synonymous mutations are known^{16,20,22–24} to affect fitness by more than 1%. Some reports have demonstrated comparable fitness effects of synonymous and nonsynonymous mutations^{25–27}. These reports, however, were based on either relatively few

genes and mutations²⁵ or many natural polymorphisms^{26,27} that may not represent random mutations. Here we test the near neutrality of synonymous mutations by measuring the fitness effects of thousands of coding mutations in 21 genes in the budding yeast *Saccharomyces cerevisiae*.

Quantifying mutational fitness effects

The 21 chosen genes participate in diverse biological processes such as metabolism, chromatin remodelling, transcription, translation, and cell wall synthesis (Supplementary Data 1) and vary by 1,000 times in their expression levels (Fig. 1a). These genes are nonessential but their deletions lower the fitness by discernable amounts²⁸ such that the mutational fitness effects are quantifiable. In each gene, we selected an approximately 150-nucleotide coding sequence and chemically synthesized all 450 possible variants that deviate from the wild type by a point mutation (Fig. 1b). The wild-type sequence at its native genomic location was replaced by the variant sequences using CRISPR–Cas9 genome editing of a haploid strain, followed by confirmation of the respiratory function of the mutant library (Extended Data Fig. 1a,b). All mutants of a gene, together with a wild-type control that went through the same CRISPR–Cas9 editing (Extended Data Fig. 1c), were

¹Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA. ²Department of Biology, Stanford University, Stanford, CA, USA. ³Present address: Microsoft, Redmond, WA, USA. ✉e-mail: jianzhi@umich.edu

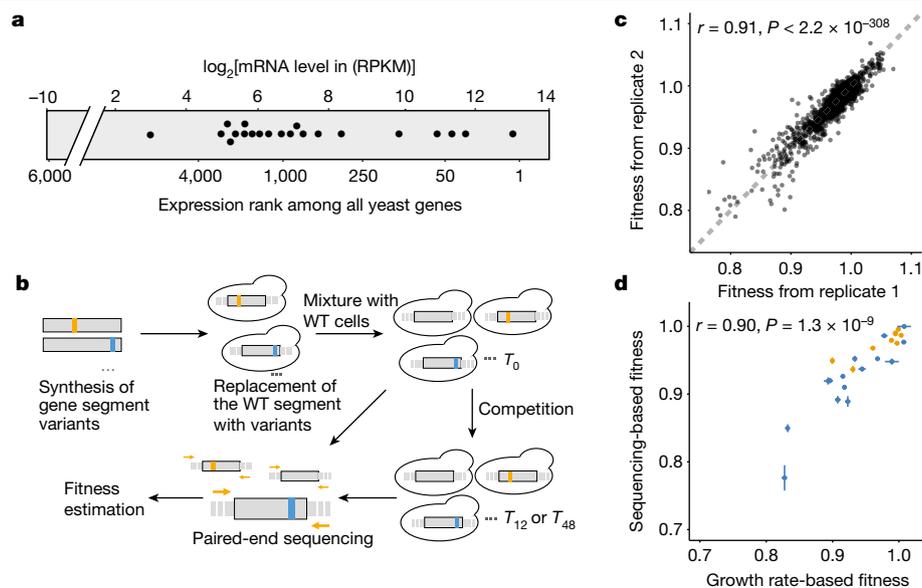


Fig. 1 | Estimating the fitness effects of coding mutations in 21 yeast genes.

a, The mRNA expression levels in YPD of the 21 genes (dots) measured in reads per kilobase of transcript per million mapped reads (RPKM) and their ranks among all yeast genes. **b**, Experimental procedure. WT, wild type. T_0 , T_{12} and T_{48} refer to the population 0, 12 and 48 h after competition, respectively. **c**, Mutant fitness estimated in the first two of four biological replicates. Each dot is a mutant ($n = 8,341$ mutants) and the dotted line indicates the diagonal. Pearson's correlation (r) and its associated P -value are presented.

d, Sequencing-based and growth rate-based fitness estimates are highly correlated. Each dot represents a synonymous (yellow) or nonsynonymous (blue) mutant. Mutants used in monoculture growth rate-based fitness estimation and those used in en masse competition followed by sequencing-based fitness estimation are constructed independently. Error bars show the standard error of the mean. Pearson's correlation r and its associated P -value are presented ($r = 0.89$ and 0.90 for the 9 synonymous and 15 nonsynonymous mutants, respectively).

competed en masse in a rich medium (YPD) at 30 °C, with no diploidization observed (Extended Data Fig. 1d). Four separate competitions were performed using a common starting population (T_0), and the focal gene was respectively amplified from T_0 and the four replicate populations at 12 (T_{12}) and 48 (T_{48}) hours, followed by 250-nucleotide paired-end Illumina sequencing (Fig. 1b). The sequences informed genotypes and enabled tabulation of genotype frequencies (Supplementary Data 2) in each population²⁹.

For the 21 genes, we identified a total of 8,341 variants with read counts of at least 50 at T_0 , including 1,866 synonymous, 6,306 nonsynonymous and 169 nonsense mutants. The observed relative numbers of synonymous and nonsynonymous mutants reflect those designed (Extended Data Fig. 2a). Changes in genotype frequencies between T_0 and T_{48} (or T_{12}) were used to estimate the fitness of each mutant relative to the wild type. The fitness estimates (Supplementary Data 3) were highly correlated between replicates, with a mean Pearson's r of 0.92 (Fig. 1c and Extended Data Fig. 2b–f). Fitness estimates from the en masse competitions agreed well with those measured from monoculture growths for 24 reconstructed synonymous and nonsynonymous mutants (Fig. 1d).

Comparing mutational fitness effects

The median fitness of the 169 nonsense mutants is 0.940 (Extended Data Fig. 3a). As expected, the corresponding value for the 6,306 nonsynonymous mutants is much higher, reaching 0.988 (Fig. 2a). Notably, the median fitness of the 1,866 synonymous mutants is 0.989, much closer to that of nonsynonymous mutants than to the neutral expectation of 1; the same trend holds for mean fitness (Fig. 2a). Whereas the fitness distribution for synonymous mutants resembles that for nonsynonymous mutants (Fig. 2a), the two distributions are statistically distinct owing to a higher density of nonsynonymous mutants in the fitness range 0.91–0.97 but a higher density of synonymous mutants in the range 0.97–0.99 (Fig. 2b and Extended Data Fig. 3b). A significant

fitness difference was observed between synonymous and nonsynonymous mutants in only five of the 21 genes, with all five exhibiting higher fitness for synonymous than nonsynonymous mutants (Fig. 2c and Extended Data Fig. 3c). Even among these five genes, however, the median fitness of synonymous mutants is much closer to that of nonsynonymous mutants than to 1 (Fig. 2c).

Classifying each mutant into one of three bins on the basis of whether its fitness was significantly below 1 (nominal $P < 0.05$, t -test), above 1 or neither, we found similar distributions for synonymous and nonsynonymous mutants (Fig. 2d). Among synonymous mutations, 75.9% are significantly deleterious, whereas 1.3% are significantly beneficial. The corresponding values are 75.8% and 1.6% for nonsynonymous mutations. Slightly lower values were obtained at the false discovery rate (FDR) of 0.05 (Fig. 2d). The smallest significant absolute fitness effect in our study is 0.001, orders of magnitude greater than the sensitivity (10^{-7}) of natural selection in yeast³⁰ (Methods). Thus, all mutations with significant fitness effects are strongly non-neutral. Mutant fitness is lower when the mutation is unobserved in the genomes of related yeast species than when it is observed (Fig. 2e and Extended Data Fig. 3d), indicating that our laboratory fitness estimates are evolutionarily relevant.

Mechanisms of mutational fitness effects

Because synonymous codon usage bias is stronger in more highly expressed genes—probably owing to translational selection⁶—synonymous mutations from the wild type are thought to be more deleterious in more highly expressed genes⁸. However, we did not detect a significant negative correlation between the expression level of a gene and the mean fitness of its synonymous mutants (Extended Data Fig. 4a). Because synonymous mutations in a gene can alter levels of its mRNA transcript^{10,12,13}, which could affect fitness³¹, we measured the relative expression level (REL) for each mutant in four replicates. In brief, from a population of cells including the wild type and all mutants of a gene, we

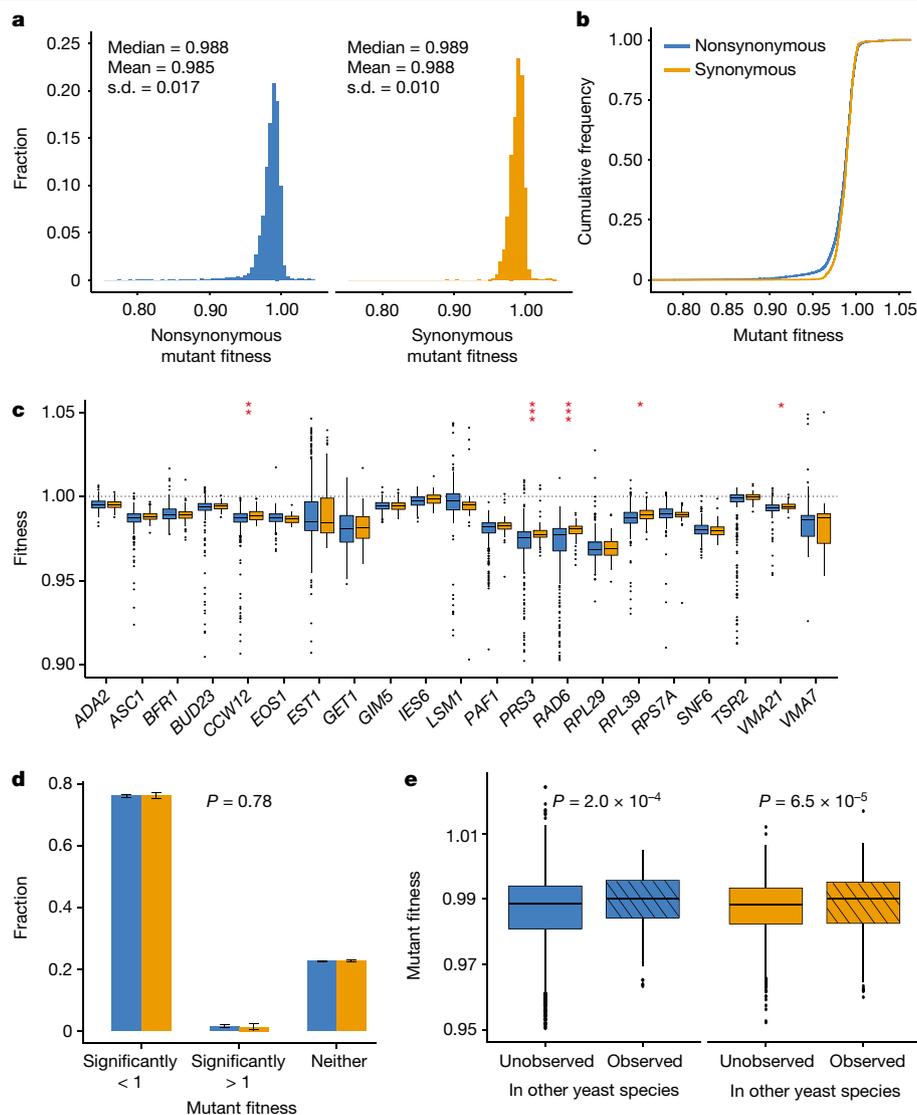


Fig. 2 | Mutant fitness in YPD. a, Distributions of the fitness of 6,306 nonsynonymous (blue) and 1,866 synonymous (yellow) mutants. The two distributions are significantly different ($P = 6.1 \times 10^{-5}$, two-tailed Wilcoxon rank-sum test; $P = 1.3 \times 10^{-6}$, Kolmogorov–Smirnov test). **b**, Cumulative frequency distributions of fitness of nonsynonymous and synonymous mutants. **c**, Fitness distributions of nonsynonymous and synonymous mutants of 21 individual genes shown by box plots. Nonsynonymous and synonymous distributions of each gene are compared by a two-tailed Wilcoxon rank-sum test followed by FDR correction ($*P < 0.05$; $**P < 0.01$, $***P < 0.001$). Mutants with fitness less than 0.9 are not shown (see Extended Data Fig. 3c for the complete figure). **d**, Fractions of nonsynonymous and synonymous mutants with fitness significantly below 1 (nominal $P < 0.05$), significantly above 1, and neither. Error bars show one standard error. Nonsynonymous and synonymous mutants are not significantly differentially distributed among the three bins

(two-tailed Fisher’s exact test). Under FDR = 0.05, 72.7% and 1.5% of nonsynonymous mutations are significantly deleterious and beneficial, respectively. The corresponding values are 72.5% and 1.1% for synonymous mutations. **e**, Mutant fitness is lower when the mutation is not observed than when it is observed in the genomes of five related yeast species. There are 5,839, 169, 1,087 and 714 mutants in the four bins (left to right). P -values are from two-tailed Wilcoxon rank-sum tests. Mutants with fitness less than 0.95 or above 1.025 are not shown (see Extended Data Fig. 3d for the complete figure). In **c** and **e**, each data point is one mutant. The lower and upper edges of a box represent the first and third quartiles, respectively, the horizontal line inside the box indicates the median, the whiskers extend to the most extreme values inside inner fences (median $\pm 1.5 \times$ interquartile range) and the dots show outliers.

amplified and sequenced the DNA of the focal gene as well as the cDNA made from the mRNA of the focal gene (Fig. 3a). REL is the number of cDNA-derived sequencing reads divided by the number of DNA-derived reads for a mutant relative to that for the wild type.

We obtained mutant RELs for 20 of the 21 genes (Supplementary Data 4). Mutant RELs are highly correlated between replicates (Extended Data Fig. 4b–g), confirming the quality of the expression estimates. REL deviates significantly from 1 (nominal $P < 0.05$, t -test) in 53.8% of synonymous and 55.0% of nonsynonymous mutants (39.7% and 39.6% at FDR < 0.05 , respectively), indicating that both synonymous and

nonsynonymous mutations frequently alter the mRNA level. The REL distribution is not significantly different between synonymous and nonsynonymous mutants (Extended Data Fig. 4h; see Extended Data Fig. 4i for individual genes) and is more or less symmetrical around 1 (Fig. 3b). By contrast, the mean REL is 0.301 for nonsense mutants (Extended Data Fig. 4j), probably owing to nonsense-mediated mRNA decay³².

Because reducing REL from 1 to 0—equivalent to gene deletion—has different fitness effects for different genes¹⁵, we rescaled mutant fitness F to $f = (F - F_0)/(1 - F_0)$, where F_0 is the fitness of the strain lacking the

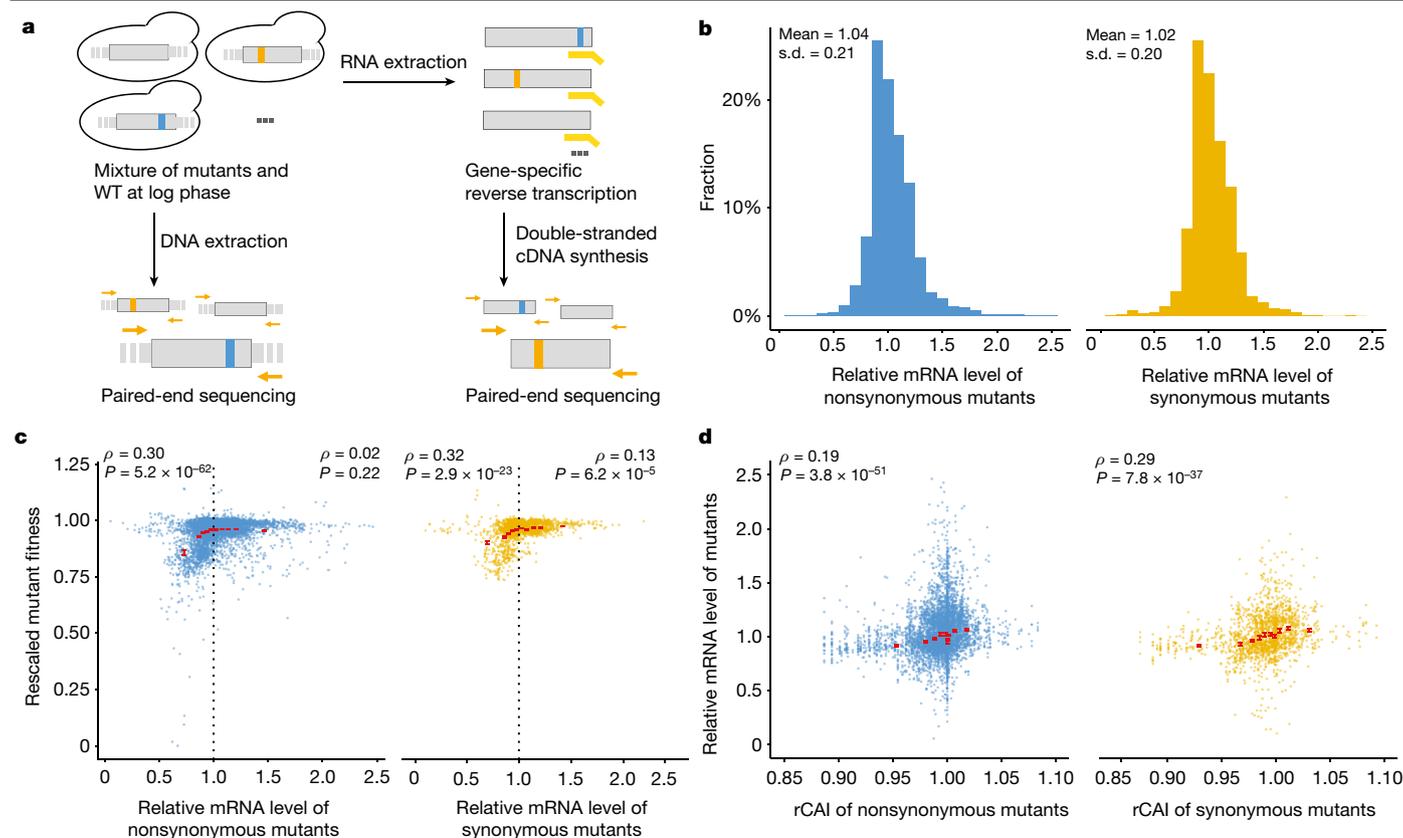


Fig. 3 | Coding mutations alter the mRNA level of the mutated gene.

a, High-throughput quantification of the mRNA levels of a focal gene in all mutants of the gene. REL, the mRNA level in a mutant relative to that in the wild type, is estimated from the number of cDNA-derived sequencing reads divided by the number of DNA-derived reads for the mutant relative to that for the WT. **b**, Frequency distributions of REL for 5,927 nonsynonymous (blue) and 1,783 synonymous (yellow) mutants. The two distributions are not significantly different ($P = 0.11$, two-tailed Wilcoxon rank-sum test). **c**, Correlation between

REL and rescaled fitness among mutants. The correlation is significantly different for mutants with REL < 1 and those with REL > 1 ($P < 0.0001$ for both nonsynonymous and synonymous mutants based on z-test after Fisher's *r*-to-*z* transformation). **d**, Positive correlation between rCAI, the CAI of a mutant relative to that of the wild type, and REL among mutants. For visualization in **c** and **d**, we group all mutants into ten bins of equal size according to their *x* values and present the mean *x* and *y* values for each bin (red dot) and the standard error of the mean *y* value (error bar).

focal gene. Consequently, $1 - f$ measures the fitness effect of a mutation relative to that of deleting the focal gene, permitting analysis of the relationship between REL and fitness across mutants of different genes. REL and rescaled fitness are significantly positively correlated for both synonymous and nonsynonymous mutants when REL is less than 1, but the correlation is weakened when REL is greater than 1 (Fig. 3c). These observations suggest that influencing the mRNA level is probably a general mechanism underlying the fitness effects of coding mutations and that reduction of expression from the wild-type level typically imposes a stronger fitness effect than increasing the expression (see Methods).

To understand how coding mutations affect the mRNA level, we identified transcription factor-binding sites in the mutated region of each gene³³, but mutations within and outside transcription factor-binding sites do not show significantly different magnitudes of expression effects (Extended Data Fig. 5a,b).

Previous manipulative experiments showed that increasing the codon adaptation index³⁴ (CAI) of a gene through synonymous mutations can boost its mRNA level by slowing mRNA degradation^{12,13,35} and perhaps enhancing transcription¹⁰. Because nonsynonymous mutations can also alter CAI, we computed the relative CAI (rCAI) of each mutant gene by dividing its CAI by that of the wild type. Indeed, a significant positive correlation exists between rCAI and REL among synonymous mutants as well as among nonsynonymous mutants (Fig. 3d). The same is true between rCAI and rescaled fitness, especially when rCAI is less than 1 (Extended Data Fig. 5c,d).

Owing to the increased prevalence of preferred codons in more highly expressed genes⁶, synonymous mutations that decrease the CAI (Extended Data Fig. 5e) or lower the mRNA level (Extended Data Fig. 5f) are more abundant in more highly expressed genes. Similar trends are seen for nonsynonymous mutations (Extended Data Fig. 5g,h), because a random nonsynonymous mutation from a preferred codon of an amino acid will probably arrive at a less preferred codon of another amino acid. Consequently, synonymous (Extended Data Fig. 5i) and nonsynonymous (Extended Data Fig. 5j) mutants of more highly expressed genes have lower mean rescaled fitness.

Because of the demand for mRNA folding strength¹¹ (MFS) which is at least in part related to translational accuracy³⁶ and co-translational protein folding³⁷, a change in MFS caused by a coding mutation may affect fitness²⁵. Indeed, we found a significant positive correlation between the relative MFS of a mutant and its rescaled fitness among mutants with reduced MFS (Extended Data Fig. 5k,l), although the correlation is substantially weaker than that between REL and rescaled fitness (Fig. 3c), suggesting that fitness effects of coding mutations are probably conferred more by their influence on the mRNA level than on the mRNA folding strength.

Fitness effects across environments

Interspecific comparisons have shown that the nonsynonymous to synonymous substitution rate ratio (d_n/d_s) is substantially below 1 for most genes in almost all organisms³⁻⁵ including yeast³⁸, indicating that

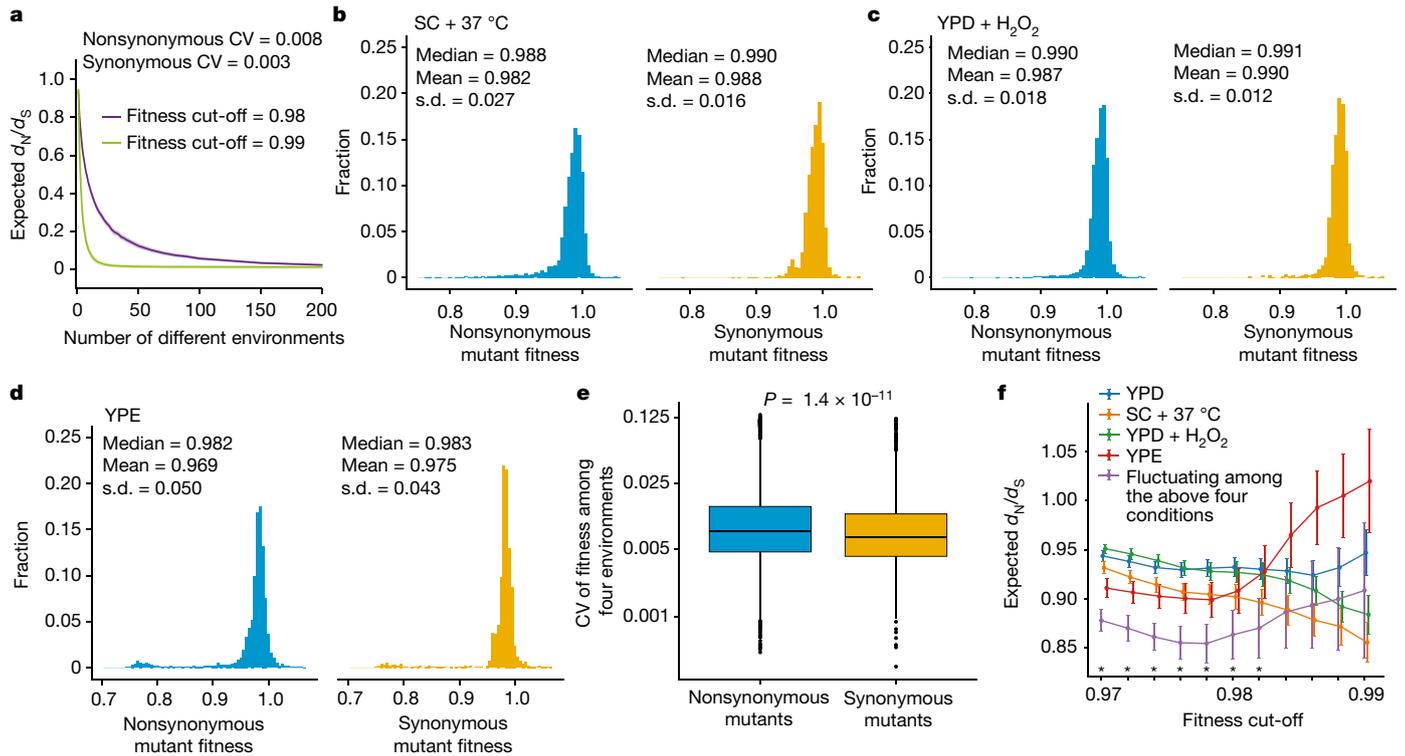


Fig. 4 | A higher fitness coefficient of variation across environments for nonsynonymous than synonymous mutants can create $d_N/d_S < 1$ despite similar DFEs of synonymous and nonsynonymous mutations in each environment. **a**, Expected d_N/d_S from 1000 simulations of a population that experiences multiple different environments. A mutant is purged if its fitness is below a preset cut-off such as 0.98 or 0.99 in any environment. Shaded areas represent 95% confidence intervals. CV, coefficient of variation. **b–d**, Distributions of nonsynonymous and synonymous mutant fitness are significantly different for growth in synthetic complete medium (SC) at 37 °C (SC + 37 °C; $P = 1.8 \times 10^{-12}$, two-tailed Wilcoxon rank-sum test; $P = 1.5 \times 10^{-9}$, Kolmogorov–Smirnov test) (**b**), YPD medium with 0.375 mM H_2O_2 ($P = 1.9 \times 10^{-7}$, two-tailed Wilcoxon rank-sum test; $P = 7.0 \times 10^{-8}$, Kolmogorov–Smirnov test) (**c**) and YPE ($P = 9.9 \times 10^{-5}$, two-tailed Wilcoxon rank-sum test; and 2.9×10^{-9} , Kolmogorov–Smirnov test) (**d**). **e**, Box plots showing distributions of fitness coefficient of variation across the four environments for 5,671 nonsynonymous

and 1,696 synonymous mutants. Box plots follow those in Fig. 2e. The mean coefficient of variation is 0.0163 for nonsynonymous and 0.0124 for synonymous mutants. The two distributions are significantly different (two-tailed Wilcoxon rank-sum test). **f**, Expected d_N/d_S when the population stays in a constant environment or a changing environment. Actual DFEs in the four individual environments are used and various fitness cut-offs are considered as in **a**. Fitness measurement error is considered through 1,000 random samples of error per mutant. The mean expected d_N/d_S and the 95% confidence interval of the expected d_N/d_S are presented. Dots and error bars are slightly shifted horizontally to help visualization. An asterisk indicates that d_N/d_S is significantly lower in the fifth population, whose environment fluctuates among the four conditions, than in each of the four constant-environment populations ($P < 0.05$). For the cut-offs where no asterisk is shown, d_N/d_S is not significantly different between the fifth population and the constant-environment population with the lowest d_N/d_S .

the probability of fixation of nonsynonymous mutations is generally much lower than that of synonymous mutations in long-term evolution, seemingly at odds with their similar distributions of fitness effects (DFEs) observed here. One possible explanation is that the two DFEs are highly dissimilar in the range of absolute fitness effects undetectable by our method, which is generally below 5×10^{-3} . For example, when beneficial mutations are ignored as in the neutral theory³⁹, if the fraction of nonsynonymous mutations with deleterious fitness effects smaller than the sensitivity of natural selection in yeast (10^{-7}) is 10% of the corresponding fraction of synonymous mutations, a d_N/d_S of approximately 0.1 will result. This hypothesis is, however, difficult to test because of the much lower sensitivity of experiments than natural selection.

We explored whether the low d_N/d_S could also be caused by a difference between synonymous and nonsynonymous mutants in their fitness variation among environments^{40,41}. Considering this variation is relevant, because the fixation of a neutral mutation takes $4N_e$ generations⁴² on average, a period during which the environment is highly likely to have changed many times. In addition to influencing the mRNA level and/or mRNA folding strength that can exert a fitness effect, nonsynonymous mutations also alter the protein sequence and potentially alter function, which synonymous mutations do

not. Because each of the molecular phenotypic effects could be environment-dependent, nonsynonymous mutants may naturally have a larger across-environment fitness variance than synonymous mutants, especially given recent reports that amino acid substitutions often show environment-specific fitness effects^{43–45}. Under the most extreme scenario, the fraction of deleterious mutations is identical between synonymous and nonsynonymous mutations in each environment, but the specific deleterious mutations vary across environments for nonsynonymous but not synonymous mutations. Consequently, when the environment of a population fluctuates within the typical fixation time, some synonymous mutations are never deleterious so may be fixed, whereas nearly every nonsynonymous mutation is deleterious under some environments so cannot be fixed, resulting in d_N/d_S much lower than 1. We investigated this model quantitatively using computer simulation. Assuming the YPD-based DFEs in each environment, we varied the fitness of a mutant among environments with a coefficient of variation greater for nonsynonymous than synonymous mutants. A mutant is selectively purged if its fitness is lower than a preset cut-off (for example, 0.99—given the fitness estimation error in our experiments) in any environment, and d_N/d_S is inferred from the fraction of unpurged nonsynonymous mutants relative to that of unpurged synonymous mutants. As predicted, d_N/d_S drops precipitously with

the number of different environments experienced by the population (Fig. 4a and Extended Data Fig. 6).

To verify the key assumption on coefficient of variation in the above model, we measured the DFEs of the same yeast synonymous and nonsynonymous mutations in three additional environments that differ in nutrient and stress, with three biological replicates per environment (Extended Data Figs. 7 and 8a–i and Supplementary Data 3). As in YPD, in each of these three environments, the median fitness of synonymous mutants is much closer to that of nonsynonymous mutants than to 1 (Fig. 4b–d) and 52.9–62.2% of synonymous mutants are significantly non-neutral (Extended Data Fig. 8j–l). These fractions are lower than that in YPD, probably because of the reduced sensitivity of our fitness measurement caused by the use of fewer replicates (Methods). For each mutant, we computed its coefficient of variation in fitness across the four environments. Indeed, the coefficient of variation is significantly greater for nonsynonymous than synonymous mutants with ($P < 10^{-5}$) or without (Fig. 4e) the control of the mean fitness in the four environments (Methods). Additionally, the fraction of neutral mutations in one environment that become deleterious in any of the other three environments is greater for nonsynonymous than for synonymous mutations (Extended Data Fig. 9). We then used the empirical DFEs and fitness estimation errors in the four environments to estimate the expected d_N/d_S after purging mutants whose fitness was lower than a cut-off in any of the environments. Indeed, comparing the four populations staying in one of the four constant environments with the fifth population whose environment fluctuates among the four conditions (Methods), we found that, in terms of d_N/d_S , the fifth population is either significantly lower than or is not statistically distinguishable from the lowest of the first four (Fig. 4f). It is expected from the simulation result (Fig. 4a) that d_N/d_S in the fifth population will further decline as the number of different environments experienced increases.

Discussion

Our characterization of the DFE of thousands of coding mutations in diverse yeast genes under four environments showed that, under any environment, most synonymous mutations are strongly non-neutral and that the DFEs of synonymous and nonsynonymous mutations are overall similar. There is no particular reason why our results would not generalize to other organisms, but confirmation experiments in diverse organisms are required to verify the generality of our findings. Because our experiments were performed in haploids, future studies should assess whether synonymous and nonsynonymous mutations also have similar DFEs in the heterozygous state.

Our results suggest a general mechanism through which coding mutations affect fitness, disturbing the mRNA level of the mutated gene, but do not preclude other mechanisms such as mRNA folding and translation. It is currently difficult to demonstrate and quantify the causal contributions of a coding mutation's various molecular phenotypic effects to its fitness effect, because this would require the difficult experiment of mimicking each molecular phenotypic effect of a coding mutation without disturbing the cell in any other aspect that might influence fitness. For instance, to mimic the influence of coding mutations on the amount of mRNA for a gene, we could use an inducible promoter to drive gene expression and adjust the promoter activity by altering the concentration of the inducer in the medium⁴⁶, but this alteration disturbs the medium composition, which could affect fitness more than through the inducible promoter. Additionally, the induction of the promoter may influence the expression of neighbouring genes. Use of tunable degrons—short amino acid sequences that regulate protein degradation⁴⁷—is another method, but degrons may also affect fitness by altering protein function or mRNA folding, and tuning degrons could disturb the medium.

The mRNA level of a gene has a strong influence on the evolutionary rate of its protein sequence, and several mechanisms of this influence

have been demonstrated^{48,49}. Our finding that the fraction of nonsynonymous mutations reducing the mRNA level rises with the mRNA level of the gene (Extended Data Fig. 5h) and the fitness ramification of this trend (Extended Data Fig. 5j) suggest an additional mechanism (Extended Data Fig. 10).

Because many biological conclusions rely on the presumption that synonymous mutations are neutral or nearly neutral^{3–5}, its invalidation has broad implications. For example, many tests infer selection on a gene by comparing its synonymous and nonsynonymous polymorphisms and/or substitutions. Given that most synonymous mutations are deleterious, making the same inference would require assuming that synonymous and nonsynonymous mutations are subject to equal selections that are unrelated to protein sequence and function. While seemingly reasonable, this assumption may not always hold¹¹, so further empirical verifications are needed. That most synonymous mutations are strongly non-neutral means that mutation rate, pattern and mechanism inferred from synonymous polymorphisms or substitutions may have been distorted. For the same reason, N_e inferred from synonymous polymorphisms in natural populations is probably substantially underestimated, impacting evolutionary studies and certain conservation-related decisions. Similarly, synonymous substitution-based dating of evolutionary divergences may be unjustifiable in some cases. Our results also imply that synonymous mutations are nearly as important as nonsynonymous mutations in causing disease and call for strengthened effort in predicting and identifying pathogenic synonymous mutations⁵⁰. Given that gene expression anomalies can cause disease⁵¹, our results further suggest the disturbance of the mRNA level as a potentially common disease mechanism of coding mutations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04823-w>.

- Kimura, M. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet Res* **11**, 247–269 (1968).
- King, J. L. & Jukes, T. H. Non-Darwinian evolution. *Science* **164**, 788–798 (1969).
- Nei, M. & Kumar, S. *Molecular Evolution and Phylogenetics* (Oxford Univ. Press, 2000).
- Li, W.-H. *Molecular Evolution* (Sinauer, 1997).
- Graur, D., Sater, A. K. & Cooper, T. F. *Molecular and Genome Evolution* (Sinauer, 2016).
- Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annu. Rev. Genet.* **42**, 287–299 (2008).
- Chamary, J. V., Parmley, J. L. & Hurst, L. D. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat. Rev. Genet.* **7**, 98–108 (2006).
- Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* **12**, 32–42 (2011).
- Stergachis, A. B. et al. Exonic transcription factor binding directs codon choice and affects protein evolution. *Science* **342**, 1367–1372 (2013).
- Zhou, Z. et al. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc. Natl Acad. Sci. USA* **113**, E6117–E6125 (2016).
- Park, C., Chen, X., Yang, J. R. & Zhang, J. Differential requirements for mRNA folding partially explain why highly expressed proteins evolve slowly. *Proc. Natl Acad. Sci. USA* **110**, E678–E686 (2013).
- Presnyak, V. et al. Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111–1124 (2015).
- Chen, S. et al. Codon-resolution analysis reveals a direct and context-dependent impact of individual synonymous mutations on mRNA level. *Mol. Biol. Evol.* **34**, 2944–2958 (2017).
- Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
- Qian, W., Yang, J. R., Pearson, N. M., Maclean, C. & Zhang, J. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet.* **8**, e1002603 (2012).
- Frumkin, I. et al. Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc. Natl Acad. Sci. USA* **115**, E4940–E4949 (2018).
- Akashi, H. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* **136**, 927–935 (1994).
- Sun, M. & Zhang, J. Preferred synonymous codons are translated more accurately: proteomic evidence, among-species variation, and mechanistic basis. *Sci. Adv.* (in the press).

19. Buhr, F. et al. Synonymous codons direct cotranslational folding toward different protein conformations. *Mol. Cell.* **61**, 341–351 (2016).
20. Walsh, I. M., Bowman, M. A., Soto Santarriaga, I. F., Rodriguez, A. & Clark, P. L. Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness. *Proc. Natl Acad. Sci. USA* **117**, 3528–3534 (2020).
21. Gilissen, C., Hoischen, A., Brunner, H. G. & Veltman, J. A. Disease gene identification strategies for exome sequencing. *Eur. J. Hum. Genet.* **20**, 490–497 (2012).
22. Agashe, D., Martinez-Gomez, N. C., Drummond, D. A. & Marx, C. J. Good codons, bad transcript: large reductions in gene expression and fitness arising from synonymous mutations in a key enzyme. *Mol. Biol. Evol.* **30**, 549–560 (2013).
23. Kristofich, J. et al. Synonymous mutations make dramatic contributions to fitness when growth is limited by a weak-link enzyme. *PLoS Genet.* **14**, e1007615 (2018).
24. Lebeuf-Taylor, E., McCloskey, N., Bailey, S. F., Hinz, A. & Kassen, R. The distribution of fitness effects among synonymous mutations in a gene under directional selection. *eLife* **8**, e45952 (2019).
25. Lind, P. A., Berg, O. G. & Andersson, D. I. Mutational robustness of ribosomal protein genes. *Science* **330**, 825–827 (2010).
26. Sharon, E. et al. Functional genetic variants revealed by massively parallel precise genome editing. *Cell* **175**, 544–557 (2018).
27. She, R. & Jarosz, D. F. Mapping causal variants with single-nucleotide resolution reveals biochemical drivers of phenotypic change. *Cell* **172**, 478–490 (2018).
28. Qian, W., Ma, D., Xiao, C., Wang, Z. & Zhang, J. The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell Rep.* **2**, 1399–1410 (2012).
29. Li, C., Qian, W., Maclean, M. & Zhang, J. The fitness landscape of a tRNA gene. *Science* **352**, 837–840 (2016).
30. Chen, P. & Zhang, J. Asexual experimental evolution of yeast does not curtail transposable elements. *Mol. Biol. Evol.* **38**, 2831–2842 (2021).
31. Keren, L. et al. Massively parallel interrogation of the effects of gene expression levels on fitness. *Cell* **166**, 1282–1294 (2016).
32. Chang, Y. F., Imam, J. S. & Wilkinson, M. F. The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.* **76**, 51–74 (2007).
33. Monteiro, P. T. et al. YEASTRACT+: a portal for cross-species comparative genomics of transcription regulation in yeasts. *Nucleic Acids Res.* **48**, D642–D649 (2020).
34. Sharp, P. M. & Li, W. H. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281–1295 (1987).
35. Radhakrishnan, A. et al. The DEAD-Box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell* **167**, 122–132 (2016).
36. Yang, J. R., Chen, X. & Zhang, J. Codon-by-codon modulation of translational speed and accuracy via mRNA folding. *PLoS Biol.* **12**, e1001910 (2014).
37. Faure, G., Ogorstov, A. Y., Shabalina, S. A. & Koonin, E. V. Role of mRNA structure in the control of protein folding. *Nucleic Acids Res.* **44**, 10898–10911 (2016).
38. Goncalves, P., Valerio, E., Correia, C., de Almeida, J. M. & Sampaio, J. P. Evidence for divergent evolution of growth temperature preference in sympatric *Saccharomyces* species. *PLoS ONE* **6**, e20739 (2011).
39. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
40. Lewontin, R. C. & Cohen, D. On population growth in a randomly varying environment. *Proc. Natl Acad. Sci. USA* **62**, 1056–1060 (1969).
41. Gillespie, J. H. Natural selection for within-generation variance in offspring number II. Discrete haploid models. *Genetics* **81**, 403–413 (1975).
42. Kimura, M. & Ohta, T. The average number of generations until fixation of a mutant gene in a finite population. *Genetics* **61**, 763–771 (1969).
43. Flynn, J. M. et al. Comprehensive fitness maps of Hsp90 show widespread environmental dependence. *eLife* **9**, e53810 (2020).
44. Dandage, R. et al. Differential strengths of molecular determinants guide environment specific mutational fates. *PLoS Genet.* **14**, e1007419 (2018).
45. Chen, P. & Zhang, J. Antagonistic pleiotropy conceals molecular adaptations in changing environments. *Nat. Ecol. Evol.* **4**, 461–469 (2020).
46. Azizoglu, A., Brent, R. & Rudolf, F. A precisely adjustable, variation-suppressed eukaryotic transcriptional controller to enable genetic discovery. *eLife* **10**, e69549 (2021).
47. Natsume, T. & Kanemaki, M. T. Conditional degrons for controlling protein expression at the protein level. *Annu. Rev. Genet.* **51**, 83–102 (2017).
48. Zhang, J. & Yang, J. R. Determinants of the rate of protein sequence evolution. *Nat. Rev. Genet.* **16**, 409–420 (2015).
49. Wu, Z. et al. Expression level is a major modifier of the fitness landscape of a protein coding gene. *Nat. Ecol. Evol.* **6**, 103–115 (2022).
50. Sauna, Z. E. & Kimchi-Sarfay, C. Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet.* **12**, 683–691 (2011).
51. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Data source

The mRNA expression levels of yeast genes in YPD (Fig. 1a) were from Chou et al.⁵². The fitness values of yeast gene deletion strains under YPD (Supplementary Data 1) were from Qian et al.²⁸. Yeast gene functions (Supplementary Data 1) were based on the *Saccharomyces* Genome Database (<https://www.yeastgenome.org/>).

Media

Standard media of YPD (1% yeast extract, 2% peptone, and 2% glucose), YPD + 0.375 mM H₂O₂, YPE (1% yeast extract, 2% peptone, and 2% ethanol) and YPG (1% yeast extract, 2% peptone, and 2% glycerol) were used. SC medium contained 0.017% yeast nitrogen base without amino acids, 0.5% sulfate, and 2% glucose, with the addition of appropriate SC mix or SC drop-out mix. 5-fluoroorotic acid (5-FOA) plates contained 0.017% yeast nitrogen base without amino acids, 0.5% sulfate, 2% glucose, SC mix and 0.15% 5-FOA.

Construction of yeast gene deletion strains

We had three primary considerations in choosing the genes for study. First, because a previous study of DFEs of synonymous and nonsynonymous mutations analysed only two ribosomal protein genes²⁵, we wanted to include genes with a larger array of functions to complement that study. Second, knowing that synonymous mutations' fitness effects may depend on the gene expression level⁸, we wanted to choose genes with a wide range of expression levels to gain a broad picture. Third, because our experiment involved deleting the gene of choice, we must study nonessential genes. Furthermore, the deletions must alter the fitness by detectable amounts such that the mutational fitness effects are quantifiable. The decision of using a 150-nucleotide region per gene was based on the read length of paired-end Illumina sequencing. The starting site of the 150-nucleotide region was randomly chosen in the first half of the coding sequence of a gene as long as the chosen 150 nucleotides are entirely within the coding region. Two exceptions were *RPL39* and *RPS7A*, where 147 nucleotides and 141 nucleotides were respectively studied because of these genes' short coding sequences.

For each chosen gene, we used CRISPR–Cas9 to delete from the genome of wild-type (BY4742) cells the 150-nucleotide target sequence and its 25-nucleotide downstream sequence that would be used as a primer binding site to amplify the gene (see Supplementary Data 5 for all primer sequences). In the deletion step, the wild-type sequence was replaced by a 23-nucleotide designed sequence (20-nucleotide Cas9 target sequence plus 3-nucleotide protospacer adjacent motif site) that would be used as the CRISPR–Cas9 recognition site in the mutant sequence insertion step. The deletion was then verified by Sanger sequencing.

Chemical synthesis of gene variants

For each gene, we GENEWIZ (<https://www.genewiz.com/en>) synthesized in an oligonucleotide mix format all 450 variants that each deviate from the wild type by a single point mutation (except for *RPL39* that had 441 variants and *RPS7A* that had 423 variants due to their shorter sequences). With the exception of oligonucleotides for *RPL39* and *RPS7A*, each oligonucleotide has 200 nucleotides, including the 150-nucleotide target sequence and its 25-nucleotide upstream and 25-nucleotide downstream flanking sequences. The flanking sequences would serve as primer binding sites for the amplification of the variant sequences. The guaranteed amount of each oligonucleotide was 3 nmol, more than enough as the DNA template for PCR amplification.

Construction of mutant libraries

The pool of the synthesized single-strand variant oligonucleotides of each gene was amplified from the oligonucleotide mix by PCR. High-fidelity Q5 polymerase (NEB) was used in all PCR reactions. The PCR-amplified double-stranded mutant sequences were transformed

along with a CRISPR–Cas9 plasmid (pML104-URA3)⁵³ into the strain with the wild-type gene deleted. The Cas9 protein would recognize the aforementioned 23-nucleotide sequence and cause double-stranded breaks. The variant sequences were inserted into the genome at the native genomic location of the focal gene via homologous recombination repair. For each gene, over 10,000 colonies were collected on SC minus uracil plates by washing with sterile water. The large number of colonies collected ensured the inclusion of most mutational variants of each gene. The variant cells were then counter-selected on the 5-FOA plates to get rid of the CRISPR–Cas9 plasmid. The cells were then stored in 30% glycerol at –80 °C.

Construction of the wild-type control

We amplified the wild-type *ASCI* gene from the genome of the haploid strain BY4742 by PCR and inserted it into the Δ *ASCI* cell using CRISPR–Cas9. Three colonies were picked and the insertion was confirmed by Sanger sequencing. The cells were then counter-selected on 5-FOA plates to remove the CRISPR–Cas9 plasmid. These three independently reconstituted wild-type strains (WT1, WT2 and WT3) were then stored in 30% glycerol at –80 °C.

We measured the maximum growth rate of BY4742 and each of the three reconstituted wild-type strains using Biotek Gen5 Microplate Reader. The cells were first grown overnight. About 5,000 cells were added into 0.1 ml YPD in a well of a Costar 96-well plate, which was in continuous shaking at 30 °C. Sixteen replicate growth curves were collected per strain, except that one replicate of BY4742 was contaminated so was discarded. The maximum growth rate was calculated following a previous protocol⁵⁴. The maximum growth rate was not significantly variable among the four strains (Extended Data Fig. 1c). For instance, the maximum growth rate of WT1 was not significantly different from that of WT2, WT3 or BY4742 (Extended Data Fig. 1c). WT1 was used as the wild-type control in en masse competitions and mutant fitness estimation. Our results would remain almost the same should the growth rate of WT2 or WT3 be used in mutant fitness calculation.

En masse competitions in YPD

A frozen sample of cells carrying the variants of a gene and a frozen sample of the wild-type control cells were revived at 30 °C in YPD (with shaking at 250 RPM) for 3 h. These cells were then mixed in an approximately 1:50 ratio of wild-type control cells to all mutant cells combined (that is, the population should contain about 2% wild-type control cells). Four replicate competitions were then started by dilution of this common starting population into four 14-ml Falcon tubes, each containing 6 ml YPD medium. Upon dilution, the cell density of the starting population was 1×10^5 cells per ml. The competition was performed in a shaking incubator (250 RPM) at 30 °C. Every 12 h, the cell culture was diluted to 1×10^5 cells per ml by transferring to 6 ml fresh YPD. The competition lasted for 48 h. The population aliquots at 0 (T_0), 12 (T_{12}), and 48 (T_{48}) hrs were stored in 30% glycerol at –80 °C. We performed a total of 84 competitions for the 21 genes (4×21).

Library preparation and Illumina sequencing

Genomic DNA was extracted from population aliquots (Masterpure Yeast DNA Purification Kit), followed by amplification of gene variants by PCR. One primer was targeted at the 25-nucleotide sequence immediately downstream of the mutated region while the other primer was annealed upstream of the mutated region beyond the homologous recombination repair sequence. This design ensured that only those variant sequences that were inserted at the native genomic location of the focal gene were amplified. The primers included Illumina sequencing adapter and i5/i7 index sequences. The amplicons were sequenced by 250-nucleotide paired-end Illumina sequencing (HiSeq2500). Paired reads for variant sequences were required to be identical to be counted. To ensure relative accuracy in fitness estimation, we considered only those genotypes with at least 50 read pairs in T_0 .

Sequencing-based fitness estimation

We estimated the fitness of each mutant relative to the wild-type control by $(P'_{MT}P_{WT})/(P_{MT}P'_{WT})^{1/G}$, where P_{MT} and P_{WT} are the respective frequencies of the mutant and wild-type control at the beginning of the competition, P'_{MT} and P'_{WT} are the corresponding frequencies at the end of the competition, and G is the number of generations of the wild-type control in the competition and equals 7.25 for 12 h and 29 for 48 h. In theory, the above formula works in an en masse competition under the assumption of no strain–strain interaction, as was confirmed by our computer simulation. The strong correlation between mutant fitness estimated from en masse competition and that estimated from monoculture growth (Fig. 1d) supports the assumption of no strain–strain interaction. To estimate G , we first allowed a frozen sample of wild-type control cells to revive at 30 °C in YPD at 250 RPM for 3 h. We then started a monoculture of the wild-type control at 1×10^5 cells per ml in 6 ml YPD. The growth continued for 12 h in a shaking incubator (250 RPM) at 30 °C. We then estimated G in the 12 h based on the culture's optical density change. G in 48 h is 4 times G in 12 h. Mutant fitness is estimated more accurately with longer competitions. However, if the fitness of a mutant was so low that the strain disappeared in T_{48} , we calculated the fitness using T_0 and T_{12} ; otherwise, we used T_0 and T_{48} . Note that only for 36 mutants were the fitness estimated using T_{12} instead of T_{48} . Based on four biological replicates, we used a t -test to determine if the fitness of a mutant deviates from 1 at the nominal P -value of 5%. The average standard error of the estimated mutant fitness was 0.005, considered as the mean detection limit of our fitness measurement. The absolute value of the smallest fitness effect with nominal $P < 0.05$ was 0.001. It has been estimated based on the level of synonymous polymorphism that N_e is approximately 10^7 in *S. cerevisiae*³⁰, suggesting that natural selection can detect a fitness effect of 10^{-7} or greater in yeast. However, if most synonymous mutations are deleterious, as the present study shows, the actual N_e would be greater than 10^7 and natural selection would be more sensitive than considered in this study.

Verifying the respiratory function of mutants

Cells from each mutant library were first serially diluted. Equal numbers of cells were then spread on YPD and YPG plates, where respiratory functions were respectively unneeded and needed for cell growth. We allowed cell growth for two days on YPD and three days on YPG, because of faster cell growth with glucose as the carbon source. Colonies were then counted on each plate. This experiment was repeated three times for the mutant library of each gene. BY4742 was used as a positive control in the respiratory function test. As a negative control, we simultaneously deleted *TOM6* and *TOM7* from BY4742, because *TOM6* and *TOM7* are components of the translocase of outer membrane (TOM) complex that is responsible for import of mitochondrially directed proteins and is important for respiration⁵⁵.

Quantifying ploidy after competition

One T_{48} population for each gene was randomly chosen and examined for ploidy. Approximately 10^7 cells were collected, washed with 1.5 ml water, and fixed by a gentle addition of 3.5 ml of 95% ethanol and incubation for 2 h at room temperature. Fixed cells were collected by centrifugation for 15 s at 10,000g, followed by resuspension of the pellet in 1 ml water and transfer to a 1.5-ml microcentrifuge tube. After a brief centrifugation, we re-suspended cells in 0.5 ml RNase solution (2 mg ml⁻¹ RNase A in 50 mM Tris pH 8.0, 15 mM NaCl, boiled for 15 min and then cooled to room temperature) and incubated the cells for at least 2 h at 37 °C. We then collected cells from the RNase solution by centrifugation for 15 s at 10,000g. Cells were incubated in 0.2 ml protease solution (5 mg ml⁻¹ pepsin and 4.5 μl ml⁻¹ concentrated HCl in H₂O) for 20 min at 37 °C and then collected by centrifugation. Cells were re-suspended in 0.5 ml 50 mM Tris pH 7.5, and were either stored at 4 °C for a few days or analysed immediately. For analysis, 50 μl of

cell suspension was transferred to 1 ml of 1 μM SYTOX Green staining solution. All samples were analysed using iQue Screener Plus flow cytometry. First, we used the forward scatter area and side scatter area with a clustering package to remove non-cell particles. Second, we used forward scatter area and forward scatter height to remove doublets. Third, we plotted DNA content histograms of the distribution of the amount of DNA per cell. We used haploid (BY4742) and diploid (BY4743) yeast cells as controls to determine ploidy. In each of these two control profiles, there are two peaks, respectively representing cells in the G1 and G2/M cell-cycle stages (1C and 2C DNA content for haploids and 2C and 4C for diploids).

Impact of PCR and sequencing errors

The following error analysis followed Li et al.²⁹. The error rate for Illumina sequencing is 3×10^{-4} per site per read (http://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf). Thus, due to sequencing error, a genotype is expected to lose $U = [1 - (1 - 3 \times 10^{-4})^{2 \times 150}] \times M_0$ read pairs, where M_0 is the true number of read pairs of the genotype and 150 is the sequence length considered. Because the fractional loss $U/M_0 = 0.086$ is a constant for all genotypes including the wild type in each sample, the loss of reads due to sequencing error does not affect fitness estimation. Sequencing error also causes the genotype to gain on average $V = (3 \times 10^{-4}/3)^2 \times M_1 = 10^{-8} M_1$ read pairs, where M_1 is the total number of read pairs for all neighbours of the focal genotype (that is, the genotypes that differ from the focal genotype by one nucleotide). Thus, the fractional gain of read pairs for the genotype is expected to be $V/M_0 = 10^{-8} \times M_1/M_0$, which has almost no impact on fitness estimation in our study. For instance, at T_0 , M_1/M_0 is expected to be 50 for the wild type and 11 for any mutant. Hence, the fractional gain of read pairs is less than 10^{-6} for any genotype.

We similarly estimated the impact of PCR error. Q5 DNA polymerase used in PCR has a very low error rate of 5.3×10^{-7} per nucleotide incorporated⁵⁶. The PCR used in sequencing library preparation had 25 cycles. Thus, owing to PCR error, a genotype is expected to lose $U = (5.3 \times 10^{-7} \times 150 \times 25) \times M_0$ molecules, where M_0 is the true number of DNA molecules of the genotype, 150 is the sequence length in nucleotides, and 25 is the number of PCR cycles. Because the fractional loss $U/M_0 = 0.002$ is a constant for all genotypes in each sample, the loss of molecules due to PCR error does not affect fitness estimation. PCR error also causes the genotype to gain on average $V = (5.3 \times 10^{-7} \times 25/3) \times M_1 = 4.4 \times 10^{-6} \times M_1$ molecules, where M_1 is the total number of molecules for all neighbours of the focal genotype. Thus, the fractional gain of molecules for the genotype is expected to be $V/M_0 = 4.4 \times 10^{-6} \times M_1/M_0$, which has little impact on fitness estimation in our study. As mentioned, at T_0 , M_1/M_0 is expected to be 50 for the wild type and 11 for any mutant. Hence, the fractional gain in the number of molecules is 2.2×10^{-4} for the wild type and 4.9×10^{-5} for any mutant.

Growth curve-based fitness estimation of reconstructed mutants

We used maximum growth rates estimated from monoculture growth curves to verify the mutant fitness estimated by en masse competition followed by sequencing. We chose nine synonymous mutants of *RPL29*, *RAD6* or *RPS7A* and 15 nonsynonymous mutants of *TSR2*, *RAD6*, *RPS7A* or *BUD23* with relatively large ranges of sequencing-based fitness estimates. We resynthesized these gene variants and remade the corresponding mutant strains. Using the method described earlier for measuring the growths of reconstituted wild-type strains, we measured the growth curves of each of these mutants as well as the wild-type control on the same 96-well plate, with eight replicates per strain. The relative fitness of a mutant was calculated by $F = 2^{\text{relative growth rate} - 1}$, where the relative growth rate is the maximum growth rate of the mutant divided by that of the wild-type control. The maximum growth rate was calculated following a previous protocol⁵⁴. The above formula of F is derived as follows. Let r be the mutant growth rate and R be the wild-type

Article

growth rate. Let T be the wild-type generation time. By definition, mutant fitness relative to the wild-type (per generation) is $F = e^{rT}/e^{RT}$. Hence, $\ln F = (r - R)T$. Because by definition $e^{RT} = 2$, $T = (\ln 2)/R$. Combining the above two equations yielded $\ln F = (r - R) \times (\ln 2)/R = (r/R - 1) \ln 2$. Therefore, $F = 2^{r/R - 1} = 2^{\text{relative growth rate} - 1}$. If mutant cells do not divide so that its population growth rate is 0, the mutant fitness relative to the WT is 0.5. If the mutation kills cells in addition to preventing mitosis, the mutant population growth rate is negative (that is, the population shrinks), which would lead to a mutant fitness that is lower than 0.5.

CRISPR–Cas9 could generate off-target mutations. However, the high fitness correlation (Fig. 1d) between two independently constructed sets of 24 mutants suggests that this potential off-target effect did not influence our result.

Identifying orthologues of the 21 *S. cerevisiae* genes in five other yeast species

To examine whether a mutation examined in *S. cerevisiae* is present in the genomes of other yeast species, we attempted to identify the orthologues of the 21 genes studied in our experiment in *Saccharomyces paradoxus*, *Saccharomyces mikatae*, *Saccharomyces uvarum*, *Saccharomyces castellii* and *Candida glabrata*, all of which diverged from *S. cerevisiae* after the whole-genome duplication in yeast. We retrieved genomic coding sequence (CDS) data from the NCBI genome assembly database (<https://www.ncbi.nlm.nih.gov/assembly/>) if they are available (*S. paradoxus*, *C. glabrata* and *S. castellii*); otherwise, we retrieved genomic DNA data (*S. mikatae* and *S. uvarum*) from the same database. For species with CDS data, we built a local blast library and performed tblastn using protein sequences of the 21 genes from *S. cerevisiae* as query sequences. The E-value threshold was set at 10^{-10} . If there was a full-length-query match, the matched subject was recorded as an orthologue. If the query was partially matched to the subject, the subject was inspected manually to ensure the orthologous relationship. For each species and gene, only the hit with the lowest E-value was examined to prevent the inclusion of paralogs. For species with genomic DNA data, we similarly built a local blast library and performed tblastn under the same E-value threshold. If there was a full-length-query match, the matched subject sequence was recorded as an orthologue. If the query was partially matched to the subject (probably owing to introns), the matched subject sequence was extended 100–2,000 bp upstream and downstream to ensure that it included all exons of the gene; the exact length of the extension was determined manually based on the length of the unmatched part of the query as well as genomic structure. We then used AUGUSTUS⁵⁷ to predict the coding region of the gene in the extended subject sequence, and manually inspected the sequence to ensure the orthologous relationship. We successfully identified almost all orthologues of the 21 genes in the five yeast species, except for *EST1*, for which we identified an orthologue only in *S. paradoxus*. We therefore excluded *EST1* from the downstream analysis. We also did not identify the *EOS1* orthologue in *S. castellii* and *IES6* orthologue in *S. mikatae*, but decided to include these two genes in downstream analysis except for the missing species. The orthologous coding sequences of the six yeasts were then aligned using MACSE v2⁵⁸. A mutation examined in *S. cerevisiae* is considered observed in the other yeasts if it appears in the genome of any of the other five yeasts and if no other nucleotide difference from *S. cerevisiae* exists in that genome in the codon harbouring the mutation; otherwise, it is considered unobserved.

Estimating the mRNA levels of mutated genes

A frozen sample of cells carrying the variants of a focal gene and a frozen sample of the wild-type control cells were revived at 30 °C in YPD with shaking at 250 RPM for 3 h. These cells were then mixed in an approximately 1:50 ratio of wild-type control cells to all mutant cells combined. Four replicate cultures were then started by diluting this common starting population into four 14 ml Falcon tubes, each containing 6 ml YPD medium. The cell density of the starting population

was 1×10^5 cells per ml. When the cells were in the log phase after 12 h of growth at 30 °C in a shaking incubator (250 RPM), we extracted DNA and RNA from the cell cultures (Masterpure Yeast DNA Purification Kit and RNeasy Mini Kit, respectively). The mRNA of the focal gene was reverse transcribed (SuperScript III First-Strand Synthesis System for RT–PCR) using about 20 nucleotides within the 25-nucleotide sequence immediately downstream of the variant sequence as the gene-specific primer.

We amplified the mutant gene segments by 25 cycles of PCR from genomic DNA and cDNA, respectively. The cDNA libraries of *EST1* were not successfully amplified, which may be because *EST1* has the lowest expression level among the 21 genes studied (Fig. 1a). As described earlier, one primer was targeted within the 25-nucleotide sequence downstream of the variant sequence while the other primer was upstream of the variant sequence and beyond the homologous recombination repair sequence. There were Illumina-adaptor and i5/i7 index sequences on the primers. The amplicons were subjected to 250-nucleotide paired-end Illumina sequencing (NovaSeq). Paired reads for variant gene sequences must be identical to be counted. To ensure accuracy in expression estimation, we excluded genotypes with fewer than 50 read pairs from the genomic DNA.

REL of a mutant is the number of cDNA-derived read pairs divided by the number of DNA-derived read pairs for the mutant, relative to the corresponding value of the wild-type control. We estimated the REL for 7,795 mutants with fitness estimates in YPD. With the four replicates in REL estimation, we used a t -test to determine if the REL of a mutant significantly deviates from 1 at a nominal P -value of 5%. Virtually identical results were obtained when REL was first log-transformed before the t -test.

Following the sequencing and PCR error analyses presented earlier, we estimated the impact of reverse transcription errors on REL estimation. The reverse transcriptase used is a version of M-MLV RT, with an error rate of 4×10^{-5} per nucleotide incorporated (<https://www.thermofisher.com/us/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/rt-education/reverse-transcriptase-attributes.html>). Due to reverse transcription error, a genotype is expected to lose $U = (4 \times 10^{-5} \times 150)M_0$ molecules, where M_0 is the expected number of cDNA molecules of the genotype and 150 is the sequence length. Because the fractional loss $U/M_0 = 0.006$ is a constant for all genotypes in each sample, the loss of molecules due to reverse transcription error does not affect expression estimation. Reverse transcription error also causes the genotype to gain on average $V = 4 \times 10^{-5} / 3 \times M_1 = 1.3 \times 10^{-5} \times M_1$ molecules, where M_1 is the expected total number of cDNA molecules for all neighbours of the focal genotype. Thus, the fractional gain of molecules for the genotype is expected to be $V/M_0 = 1.3 \times 10^{-5} \times M_1/M_0$, which has little impact on expression estimation in our study. M_1/M_0 is expected to be about 50 for the wild-type and 11 for mutants whose expression levels are comparable with that of the wild type. The corresponding fractional gains of molecules are 6.5×10^{-4} and 1.4×10^{-4} , respectively. Even if a mutant has a REL as low as 0.1, M_1/M_0 is 110 and the fractional gain of the number of molecules is 1.4×10^{-3} . As described, PCR and sequencing errors had virtually no effect. Thus, the overall error from reverse transcription, PCR, and sequencing is negligible in expression estimation.

In addition to correlating mutant REL with rescaled fitness (Fig. 3c), we used a linear mixed model to assess the relative importance of REL and mutation type (synonymous versus nonsynonymous) to rescaled fitness, with gene identity added as a random effect. We separately analysed mutants with REL < 1 and those with REL > 1, because of their apparently different relationships with rescaled fitness (Fig. 3c). For mutants with REL < 1, the fraction of variance of rescaled fitness explained by REL is 61.5% ($P < 2.2 \times 10^{-16}$), while that explained by mutation type is only 0.2% ($P = 0.0002$). For mutants with REL > 1, the fraction of variance of rescaled fitness explained by REL is 7.4% ($P < 2.2 \times 10^{-16}$),

while that explained by mutation type is only 0.4% ($P = 1.5 \times 10^{-7}$). These results demonstrate that REL explains a substantially larger fraction of variance of rescaled fitness than does mutation type.

Additionally, after accounting for gene-specific effects using a mixed-effect model, we found the positive correlation between the rescaled fitness and REL to remain significant when $REL < 1$ ($P = 2.3 \times 10^{-47}$). There is a marginally significant negative correlation between the rescaled fitness and REL when $REL > 1$ ($P = 0.048$). We also attempted to fit a quadratic model using $\log_2(REL)$ as an independent variable and accounted for a random effect of gene identity. Indeed, the hypothesis that the fitness peak is at $REL = 1$ could not be rejected.

Codon adaption index

We computed CAI for the entire coding sequence of each wild-type or mutant gene, using previously reported yeast relative synonymous codon usage estimates³⁴, which are highly correlated with those derived from the 200 most highly expressed genes¹⁵ ($r = 0.995$).

mRNA folding strength

The minimum free energy at 30 °C was calculated for each wild-type or mutant mRNA sequence using RNAfold in ViennaRNA (2.4.17) with default parameters except for the temperature⁵⁹. We define MFS as the absolute value of the minimum free energy.

Transcription factor-binding sites

Transcription factor-binding sites were searched in the wild type for the 150-nucleotide target sequence plus the 20-nucleotide flanking sequence on each side using the database Yeastract³³.

DFE estimation in SC + 37 °C, YPD + 0.375mM H₂O₂ and YPE

The experiment followed that in DFE estimation in YPD, except that the competitions lasted for 20 generations (cells were transferred 6.5 and 13 generations after the start of the competition) and had three replicates per environment. Sequencing library preparation was unsuccessful for mutants of *ESTI* and *PAFI*, probably because of primer degradations. Therefore, we acquired the fitness data of mutants of 19 genes in these three additional environments. The fraction of mutants whose fitness is significantly different from 1 is lower here than in YPD, probably because of the reduced statistical power due to the lowered number of replications. Indeed, when we randomly sampled three of the four replicates from YPD, the fraction of mutants whose fitness is significantly different from 1 (nominal $P < 0.05$) decreased to an average of 0.63 and 0.64 for synonymous and nonsynonymous mutants, respectively, similar to those observed in these three additional environments (Extended Data Fig. S3j–l). To examine whether the difference between synonymous and nonsynonymous mutants in fitness coefficient of variation across the four environments is entirely due to a potential difference in mean fitness, we controlled the mean fitness in the four environments when comparing the across-environment fitness coefficient of variation between synonymous and nonsynonymous mutants. Specifically, we used an identity index of 0 for each synonymous mutant and 1 for each nonsynonymous mutant. The partial Spearman's correlation between the identity index and coefficient of variation upon the control of the mean fitness in the four environments is 0.052 ($P = 7.7 \times 10^{-6}$).

Simulation of the impact of environmental changes on d_N/d_S

Our simulation assumed that the DFEs of synonymous and nonsynonymous mutations estimated from YPD hold in each environment, but the fitness effect of a mutation can vary across environments. We respectively constructed cumulative fitness distribution functions of synonymous and nonsynonymous mutants from the corresponding fitness data collected in YPD. We started from all synonymous mutants with fitness measured in YPD, and ranked these mutants from low to high by their YPD fitness. We then added a random noise drawn from

the normal distribution $N(0, \sigma^2)$ to each fitness value, and ranked the mutants by their new fitness values. Let us assume that, after the addition of noise, the mutant originally ranked i now had a rank of j . We then randomly sampled M synonymous mutants from the cumulative fitness distribution function and ranked them by their fitness, where M is the number of synonymous mutants with fitness measured in YPD. We assigned the fitness of the mutant ranked the j th in these M sampled mutants to mutant i as its fitness in a new environment. The above procedure was repeated for each environment considered. Fitness coefficient of variation among environments was controlled by adjusting σ^2 , with larger σ^2 yielding greater coefficient of variation. Many σ^2 values were tried to achieve a target coefficient of variation (difference between observed and target coefficient of variation < 0.0001). The same was done for nonsynonymous mutants. We set a higher coefficient of variation for nonsynonymous than synonymous mutants. We set a fitness cut-off (0.98 or 0.99) and assumed that any mutant with fitness below the cut-off in any environment was purged. We then computed d_N/d_S by the fraction of unpurged nonsynonymous mutants divided by the fraction of unpurged synonymous mutants. Under each parameter set, we repeated the simulation 1,000 times and reported the mean d_N/d_S and its 95% confidence interval.

Expected d_N/d_S in the four environments examined

To predict the expected d_N/d_S in long-term evolution in each of the four environments where DFEs were measured here, we considered all of the synonymous and nonsynonymous mutants with fitness measured in the environment. Because the fitness measures contained measurement errors, we added a random error term drawn from the normal distribution $N(0, \sigma_{se}^2)$ to the measured fitness, where σ_{se} is the mutant-specific standard error of the measured fitness estimated from the experimental replicates in the environment. We set a fitness cut-off and assumed that any mutant with fitness in the environment below the cut-off was purged. We then computed d_N/d_S by the fraction of unpurged nonsynonymous mutants divided by the fraction of unpurged synonymous mutants. In an environment that varies among the four individual conditions, we assumed that any mutant with fitness below the cut-off in any condition was purged. Because of random measurement errors considered, we repeated the prediction 1000 times and presented the 95% confidence interval of the predicted d_N/d_S .

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Sequencing data generated in this study have been deposited into NCBI with the Bioproject ID PRJNA750109. Public data used include gene function annotations in the *Saccharomyces* Genome Database (<https://www.yeastgenome.org/>) and genomic coding sequences of *S. paradoxus*, *C. glabrata*, and *S. castellii* and genomic sequences of *S. mikatae* and *S. uvarum* from the NCBI genome assembly database (<https://www.ncbi.nlm.nih.gov/assembly/>). Source data are provided with this paper.

Code availability

Custom code is available at <https://github.com/song88180/Mutational-Fitness-Effects> and <https://doi.org/10.5281/zenodo.5908478>.

52. Chou, H. J., Donnard, E., Gustafsson, H. T., Garber, M. & Rando, O. J. Transcriptome-wide analysis of roles for tRNA modifications in translational regulation. *Mol. Cell* **68**, 978–992 (2017).
53. Laughery, M. F. et al. New vectors for simple and streamlined CRISPR–Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast* **32**, 711–720 (2015).

Article

54. Warringer, J., Ericson, E., Fernandez, L., Nerman, O. & Blomberg, A. High-resolution yeast phenomics resolves different physiological features in the saline response. *Proc. Natl Acad. Sci. USA* **100**, 15724–15729 (2003).
55. Honlinger, A. et al. Tom7 modulates the dynamics of the mitochondrial outer membrane translocase and plays a pathway-related role in protein import. *EMBO J.* **15**, 2125–2137 (1996).
56. Potapov, V. & Ong, J. L. Examining sources of error in PCR by single-molecule sequencing. *PLoS ONE* **12**, e0169774 (2017).
57. Stanke, M. & Morgenstern, B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* **33**, W465–W467 (2005).
58. Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N. & Delsuc, F. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Mol. Biol. Evol.* **35**, 2582–2584 (2018).
59. Hofacker, I. L. et al. Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.* **125**, 167–188 (1994).
60. Zhang, J. & He, X. Significant impact of protein dispensability on the instantaneous rate of protein evolution. *Mol. Biol. Evol.* **22**, 1147–1155 (2005).

Acknowledgements We thank P. Chen, H. Liu, and H. Xu for technical assistance and W. Qian, X. Wei, J.-R. Yang, and members of the Zhang laboratory for valuable comments. This work was supported by the U.S. National Institutes of Health research grant R35GM139484 to J.Z.

Author contributions J.Z. conceived the study. X.S., C.L., and J.Z. designed the study. X.S. performed the experiments. X.S. and S.S. analysed the data. X.S. and J.Z. wrote the paper with input from all authors.

Competing interests The authors declare no competing interests.

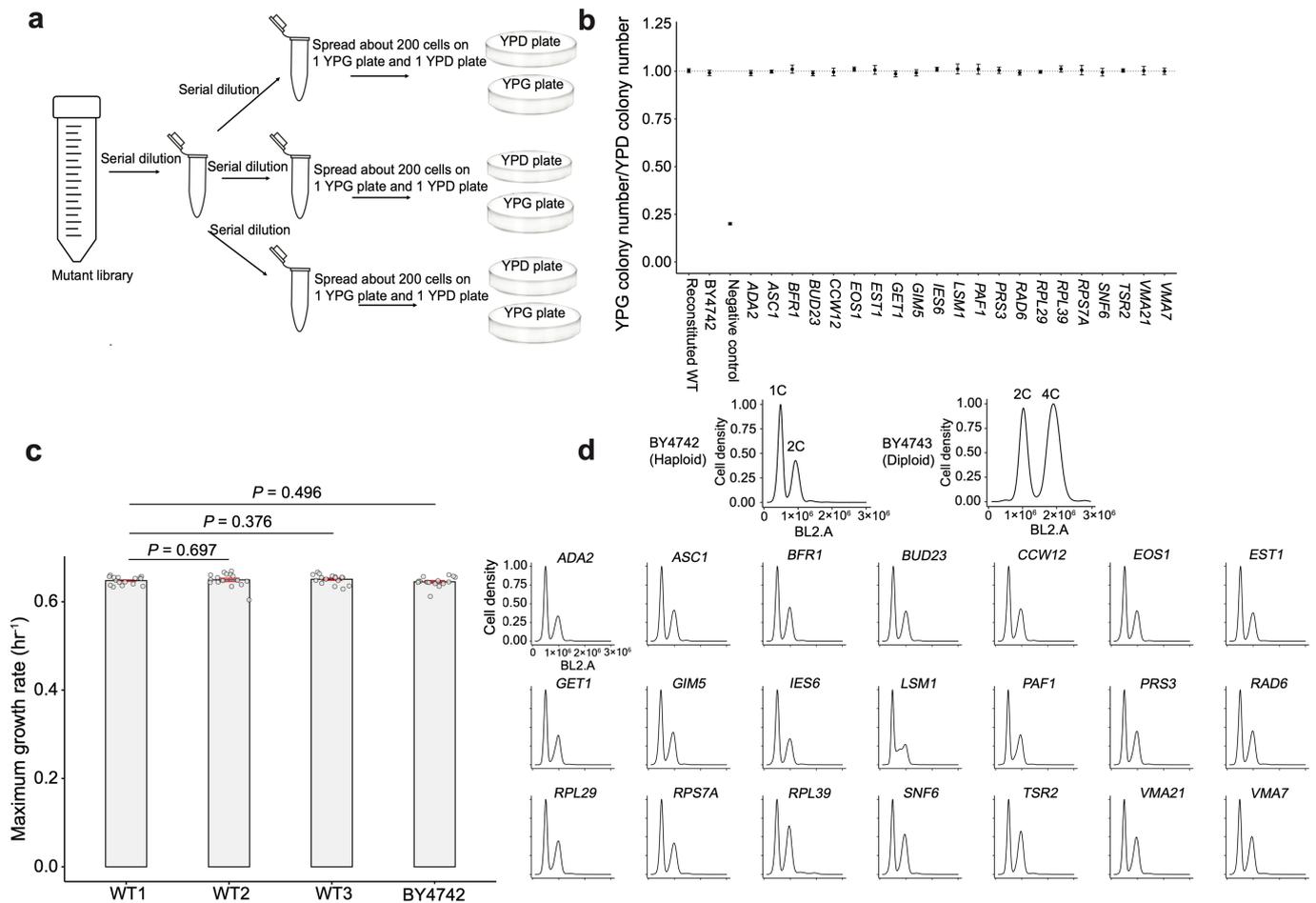
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04823-w>.

Correspondence and requests for materials should be addressed to Jianzhi Zhang.

Peer review information *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

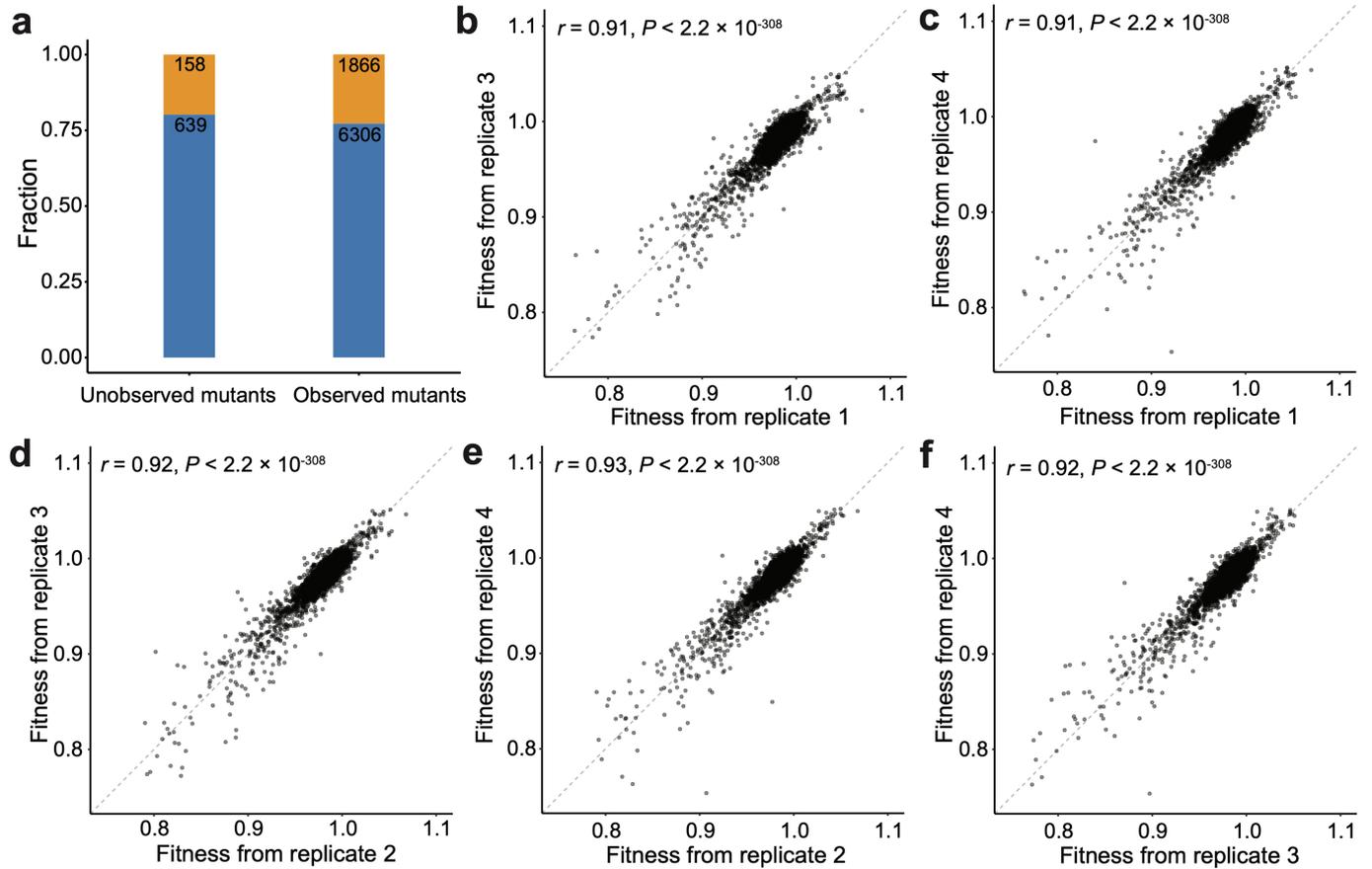
Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Properties of wild-type and mutant strains analyzed.

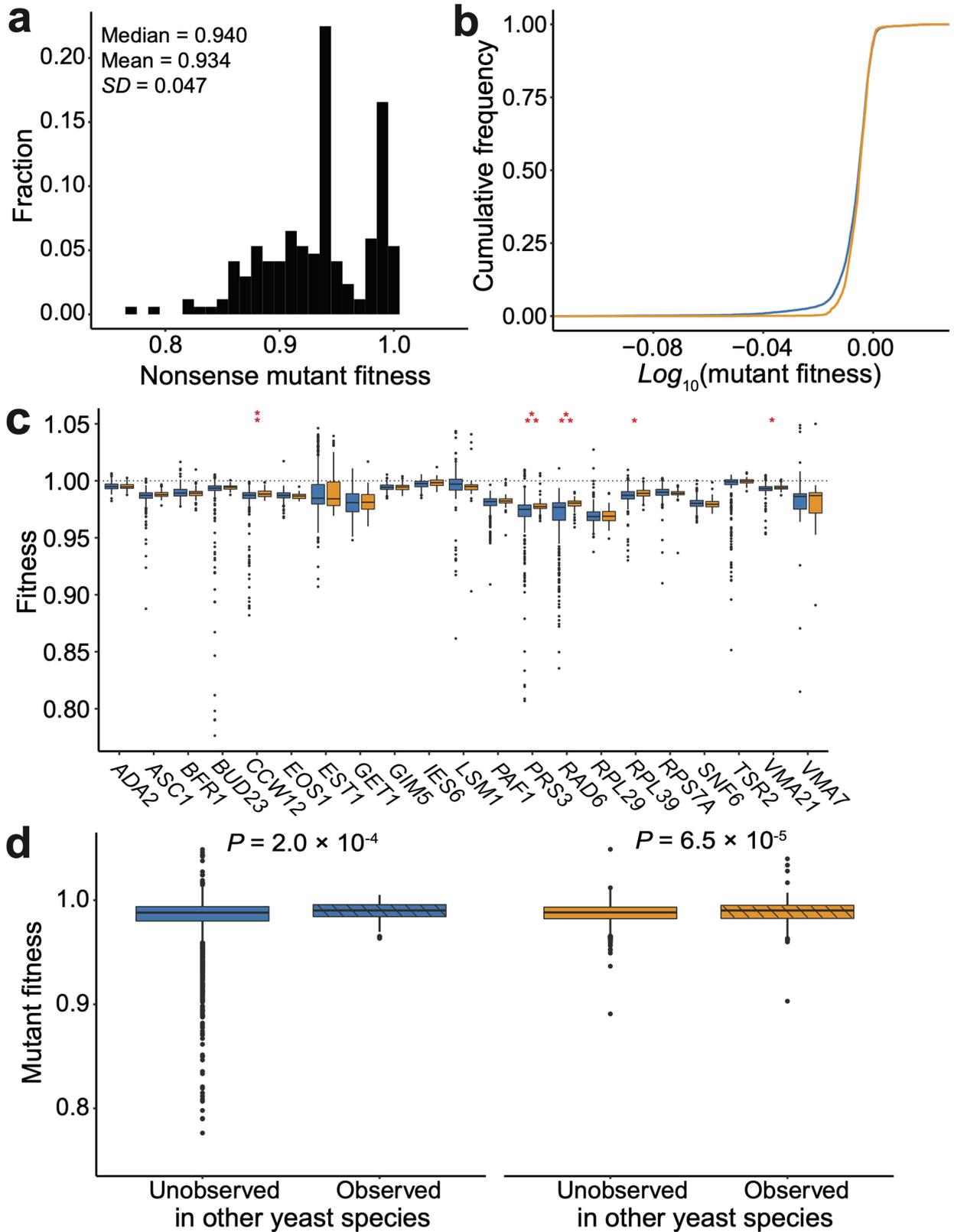
a, Experimental procedure for testing cellular respiratory functions. Cells from each of the 21 mutant libraries were spread on YPD and YPG plates, followed by colony counting after growth. Respiration is needed for cell growth on YPG but not on YPD. **b**, Mean ratio of YPD colony number to YPG colony number for each mutant library, based on three replicates per library. Error bars show the standard error of the mean. The negative control is deficient in respiration due to gene deletions (see Methods). **c**, Maximum growth rates of three reconstituted wild-type strains and BY4742. WT1 was used as the wild-type control in *en masse* competitions with mutants. The red error bar indicates the

standard error of the mean based on 16 replicates each shown by a dot (15 for BY4742). *P*-values are from two-tailed *t*-tests. The growth rate is not significantly different among the four strains ($P = 0.58$, one-factor ANOVA test). **d**, Ploidy of one T_{48} population per mutant library assessed by flow cytometry. SYTOX Green fluorescence was analyzed using the BL2 detector that measured the output from the 488-nm laser (blue). In control flow cytometry profiles, the two peaks respectively represent cells in the G1 and G2/M cell-cycle stages (1C and 2C DNA content for haploids while 2C and 4C for diploids).



Extended Data Fig. 2 | Mutant fitness quantification. **a**, Fractions of synonymous (yellow) and nonsynonymous (blue) mutants among designed but unobserved mutants and those among observed mutants. Nonsense mutants are not considered. Numbers in the bars are numbers of mutants. The distributions of synonymous and nonsynonymous mutants among the unobserved and observed mutant groups are not significantly different

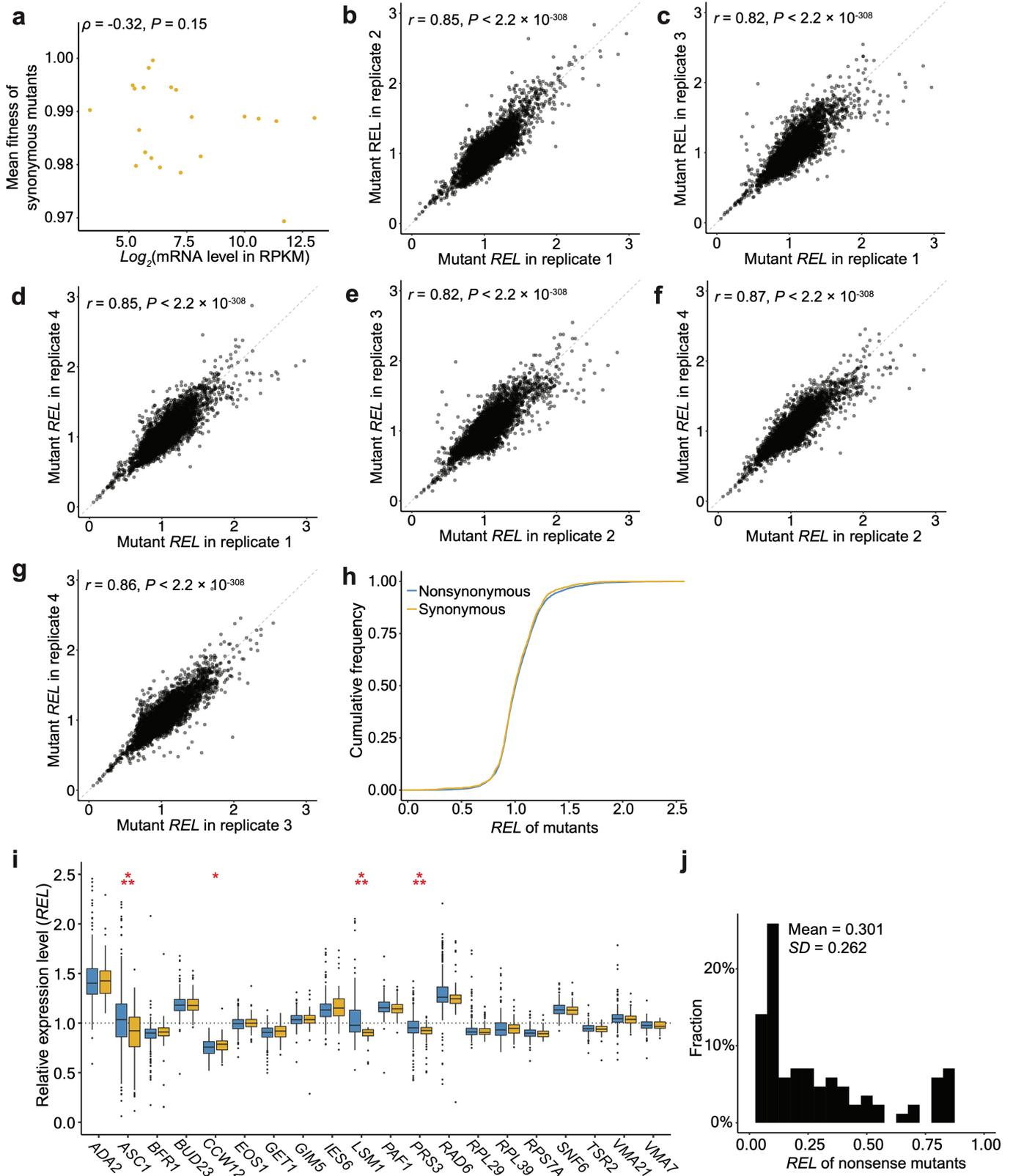
($P > 0.05$, Fisher's exact test). **b-f**, Correlation between every two of the four replicates in estimated mutant fitness under YPD at 30 °C. The correlation between replicate 1 and replicate 2 is presented in Fig. 1c. Each dot is a mutant and the dotted line indicates the diagonal. Pearson's correlation r and its associated P -value are presented. Among-genotype sum of squares explains 93.8% of the total sum of squares (one-factor ANOVA).



Extended Data Fig. 3 | Mutant fitness distribution under YPD at 30 °C.
a, Distribution of the fitness of 169 nonsense mutants. The peak around 0.94 is caused by 26 nonsense mutants of *GET1* that all have fitness of about 0.94.
b, Cumulative frequency distributions of $\log_{10}(\text{mutant fitness})$ of

nonsynonymous (blue) and synonymous (yellow) mutants. **c**, The full figure of Fig. 2c, including low-fitness mutants that are not shown in Fig. 2c. **d**, The full figure of Fig. 2e, including low-fitness and high-fitness mutants that are not shown in Fig. 2e.

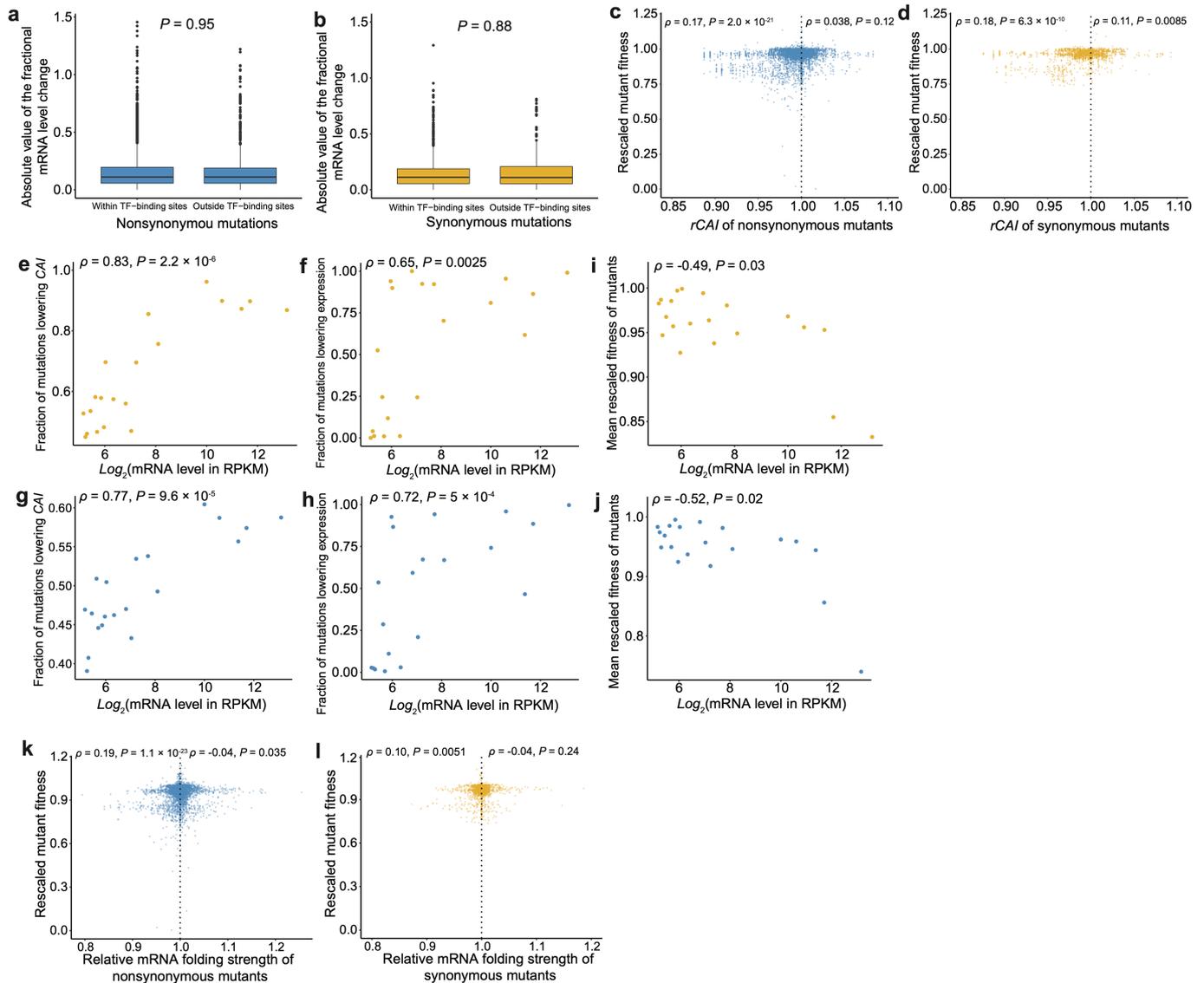
Article



Extended Data Fig. 4 | See next page for caption.

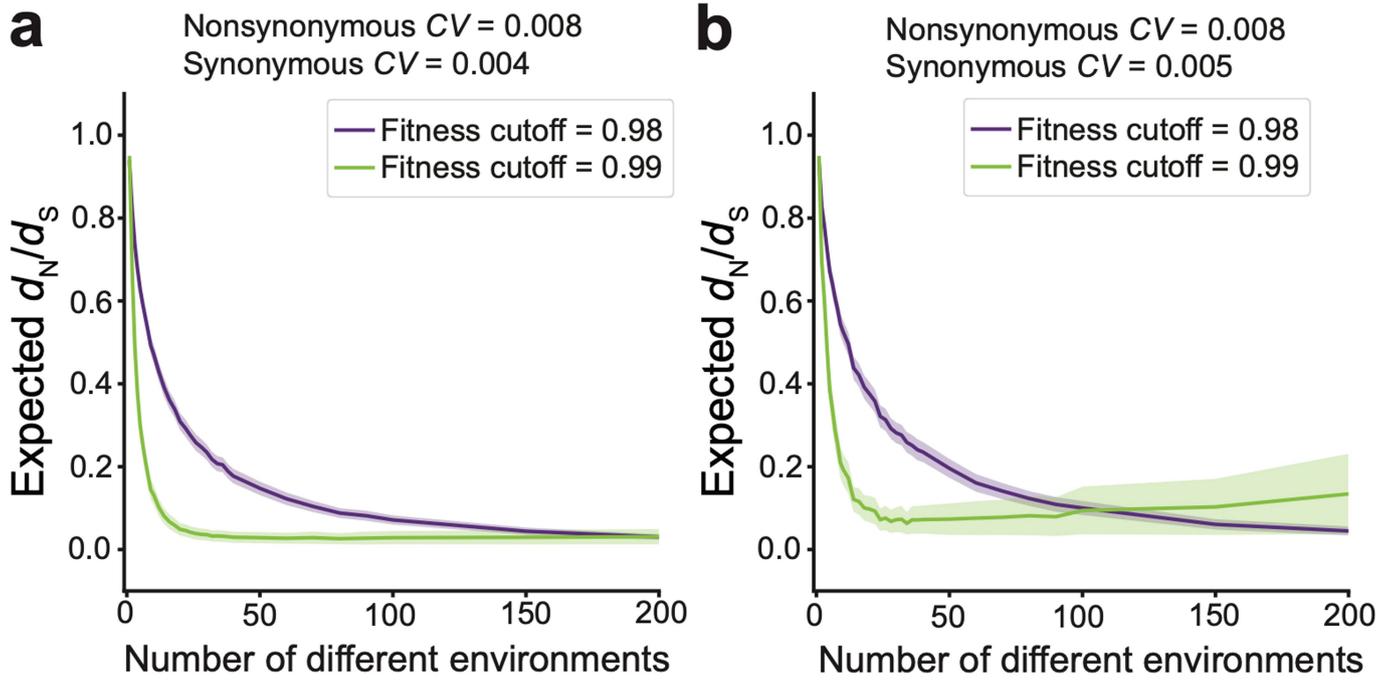
Extended Data Fig. 4 | Coding mutations influence the mRNA level of the mutated gene. **a**, Non-significant negative correlation between the mean fitness of synonymous mutants of a gene and the expression level of the gene. Each dot represents a gene. Spearman's correlation ρ and associated P -value are presented. **b–g**, Correlation in mutant REL between replicates, which are indicated on the axes of each panel. Each dot is a mutant, and the dotted line indicates the diagonal. Pearson's correlation r and its associated P -value are presented. Among-genotype sum of squares explains 89.7% of total sum of squares (one-factor ANOVA). **h**, Cumulative frequency distributions of REL of nonsynonymous and synonymous mutants. **i**, Relative expression level (REL)

distributions of nonsynonymous (blue) and synonymous (yellow) mutants of 20 individual genes shown by box plots. The lower and upper edges of a box represent the first (qu_1) and third (qu_3) quartiles, respectively, the horizontal line inside the box indicates the median (md), the whiskers extend to the most extreme values inside inner fences, $md \pm 1.5(qu_3 - qu_1)$, and the dots show outliers. Nonsynonymous and synonymous distributions of each gene are compared by a two-tailed Wilcoxon rank-sum test, with FDR-adjusted P -values indicated as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **j**, Distribution of REL of nonsense mutants.



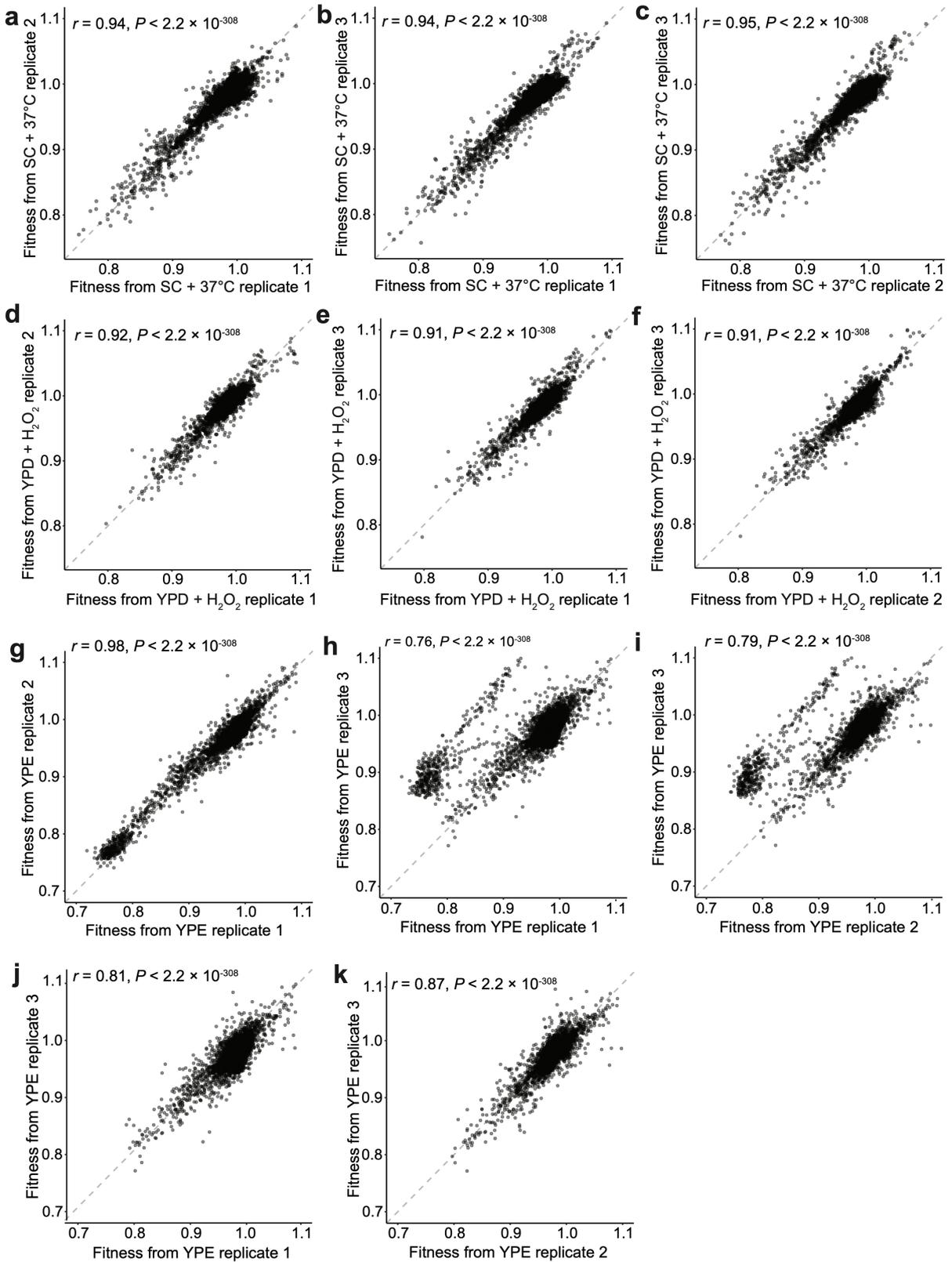
Extended Data Fig. 5 | Mechanisms underlying coding mutations' fitness effects. **a–b**, Box plots showing similar absolute fractional changes in the mRNA level induced by nonsynonymous (**a**) or synonymous (**b**) mutations within and outside TF-binding sites. The lower and upper edges of a box represent the first (qu_1) and third (qu_3) quartiles, respectively, the horizontal line inside the box indicates the median (md), the whiskers extend to the most extreme values inside inner fences, $md \pm 1.5(qu_3 - qu_1)$, and the dots show outliers. P -values are from two-tailed Wilcoxon rank-sum test ($n = 1191, 4736, 367$, and 1411 , respectively, for the four bars from left to right). **c–d**, Positive correlation between $rCAI$ and rescaled fitness among nonsynonymous (**c**) and synonymous (**d**) mutants, respectively. **e**, Fraction of synonymous mutations lowering CAI increases with the expression level of the gene. **f**, Fraction of synonymous mutations lowering the expression level increases with the expression level of the gene. **g**, Fraction of nonsynonymous mutations lowering CAI increases with the expression level of the gene. **h**, Fraction of

nonsynonymous mutations lowering the expression level increases with the expression level of the gene. **i**, Mean rescaled fitness of synonymous mutants declines with the expression level of the gene. **j**, Mean rescaled fitness of nonsynonymous mutants declines with the expression level of the gene. Because deleting a more highly expressed gene tends to cause a greater fitness reduction⁶⁰, the finding in panel **j** means that the mean fitness reduction caused by a nonsynonymous mutation should rise with the expression level of the gene. In **e–j**, each dot represents a gene. **k–l**, positive correlation between the relative mRNA folding strength ($rMFS$) of a nonsynonymous (**k**) or synonymous (**l**) mutant and its rescaled fitness when $rMFS$ is below 1. The $rMFS$ of a mutant is its mRNA folding strength (i.e., the absolute value of its minimal folding energy) divided by that of the wild-type. In each panel, the correlation is separately computed for mutants with $rMFS < 1$ and those with $rMFS > 1$. In **c–l**, rank correlations (ρ) and associated P -values are shown.



Extended Data Fig. 6 | A higher coefficient of variation (CV) of fitness across environments for nonsynonymous than synonymous mutants can create a nonsynonymous to synonymous substitution rate ratio (d_N/d_S) that is substantially below 1 despite similar fitness effects of synonymous and nonsynonymous mutations in each environment. **a, Mean expected d_N/d_S from 1000 simulations of a population that experiences multiple different environments. A mutant is purged if its fitness is lower than a preset cutoff such as 0.98 or 0.99 in any environment. Shaded areas represent 95% confidence intervals. **a**, Results with $CV = 0.004$ for synonymous mutants. **b**, Results with**

$CV = 0.005$ for synonymous mutants. Note that, under the fitness cutoff of 0.99, d_N/d_S starts to increase with the number (m) of environments when m is large. Raising m reduces the fraction of synonymous mutations that are always neutral (FAN_S) as well as the fraction of nonsynonymous mutations that are always neutral (FAN_N). Because the fitness CV is larger for nonsynonymous than synonymous mutants in the simulation, FAN_N decreases with m more quickly than does FAN_S when m is small. When m is large, FAN_N is small, making it possible for FAN_S to decrease with m more quickly than FAN_N . As a result, d_N/d_S might increase with m when m is large.



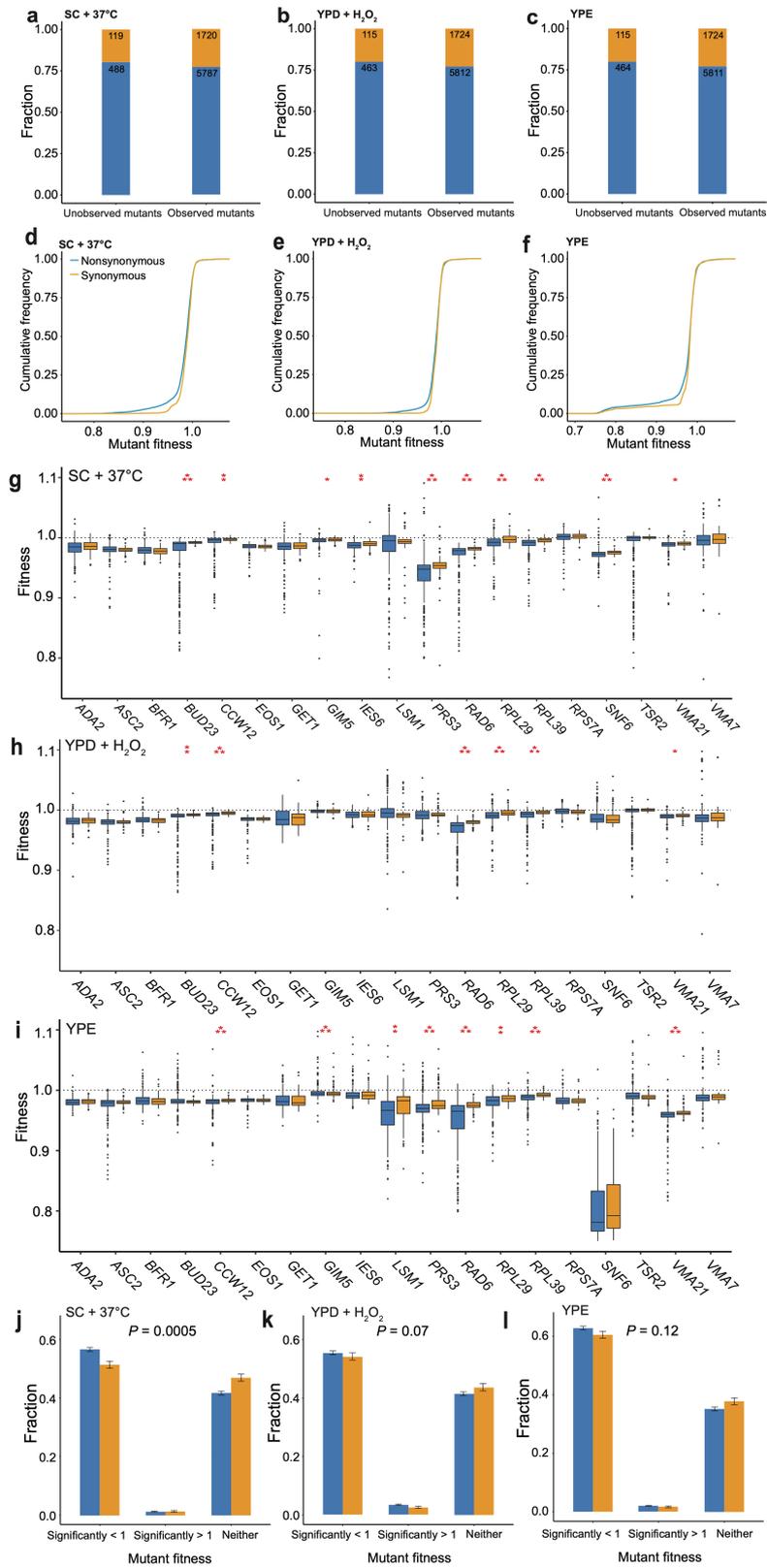
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Pairwise correlation between replicates in estimated mutant fitness in each of the three additional environments used.

a–c, Correlation between every two of the three replicates in estimated mutant fitness under SC at 37 °C. Each dot is a mutant and the dotted line indicates the diagonal. Pearson's correlation r and its associated P -value are presented.

Among-genotype sum of squares explains 96.1% of the total sum of squares (one-factor ANOVA). **d–f**, Correlation between every two of the three replicates in estimated mutant fitness under YPD + 0.375 mM H₂O₂. Among-genotype sum of squares explains 94.4% of the total sum of squares. **g–i**, Correlation between

every two of the three replicates in estimated mutant fitness under YPE. **j**, Correlation between replicates 1 and 3 in estimated mutant fitness under YPE after the exclusion of *SNF6* mutants. **k**, Correlation between replicates 2 and 3 in estimated mutant fitness under YPE after exclusion of *SNF6* mutants. Panels **g–k** suggest that the fitness estimates of *SNF6* mutants in replicate 3 under YPE are unreliable, so are unused in fitness estimation in YPE. When *SNF6* is excluded, among-genotype sum of squares explains 91.0% of the total sum of squares in YPE.

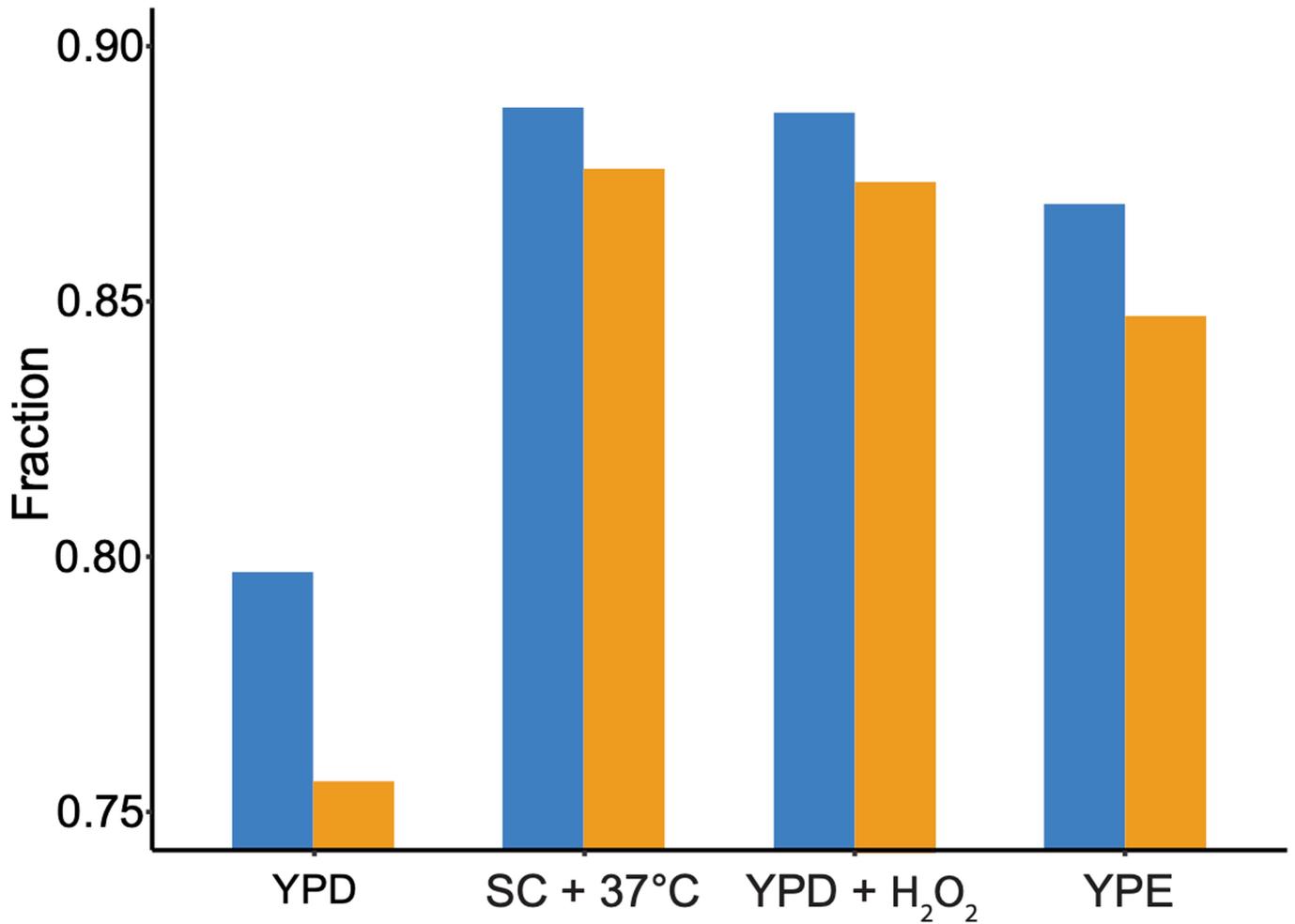


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Mutant fitness in the three additional environments used. **a–c**, Fractions of synonymous (yellow) and nonsynonymous (blue) mutants among designed but unobserved mutants and those among observed mutants in each environment. Nonsense mutants are not considered. Numbers in the bars are numbers of mutants. The distributions of synonymous and nonsynonymous mutants among the unobserved and observed mutant groups are not significantly different in each environment ($P > 0.05$, Fisher's exact test). **d–f**, Cumulative frequency distributions of fitness of nonsynonymous and synonymous mutants in each environment. **g–i**, Fitness distributions of nonsynonymous and synonymous mutants of 19 individual genes shown by box plots in each environment. The lower and upper edges of a box represent the first (qu_1) and third (qu_3) quartiles, respectively, the horizontal line inside the box indicates the median (md), the whiskers extend to the most extreme

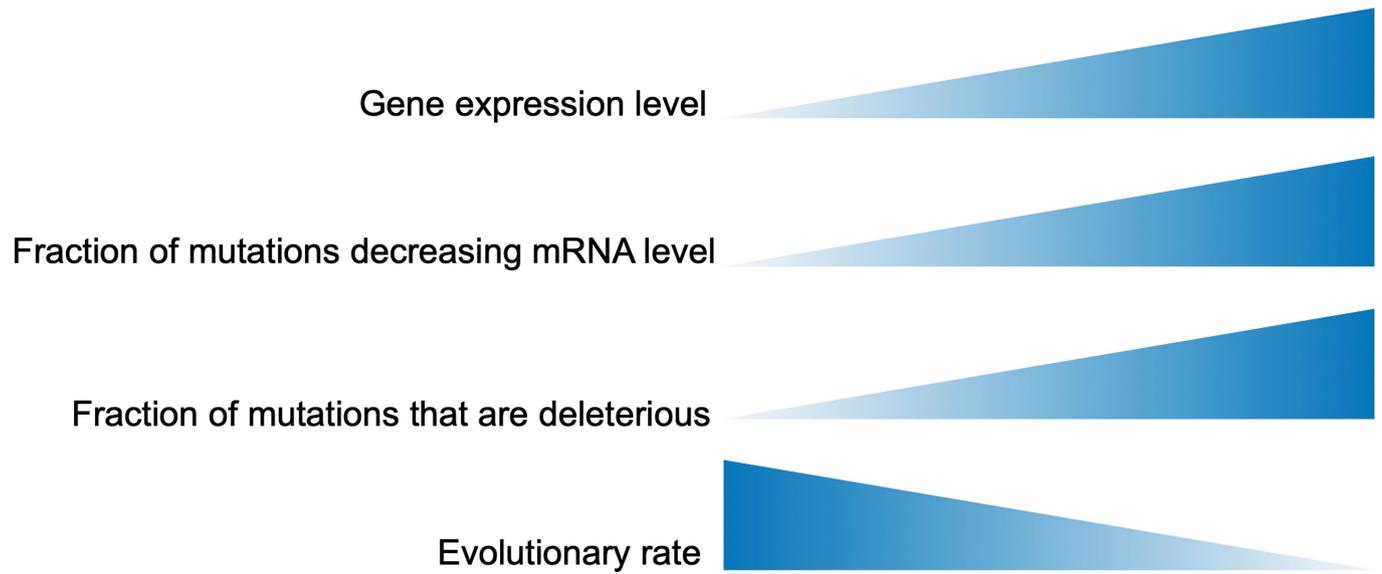
values inside inner fences, $md \pm 1.5(qu_3 - qu_1)$, and the dots show outliers. Nonsynonymous and synonymous distributions for each gene are compared by a two-tailed Wilcoxon sum-rank test, with the FDR-adjusted P -value indicated as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **j–l**, Fractions of mutants with fitness significantly below 1 ($P < 0.05$), significantly above 1, and neither, respectively, in each environment. The error bar shows one standard error. The distributional difference between synonymous and nonsynonymous mutants among the three bins is tested by two-tailed Fisher's exact test, with the P -value indicated. At FDR = 0.05, 40.7% and 0.7% of nonsynonymous mutations and 34.8% and 0.5% of synonymous mutations are significantly deleterious and beneficial, respectively, in SC+37 °C. These values become 35.5%, 1.7%, 31.9% and 1.6% in YPD+H₂O₂, and 47.6%, 1.4%, 45.6%, and 1.0% in YPE.

Article



Extended Data Fig. 9 | Fractions of nonsynonymous (blue) and synonymous (yellow) neutral mutations in one environment (indicated on the X-axis) that become deleterious in any of the other three environments. The fractions are higher for nonsynonymous than synonymous mutations

($P < 0.05$, paired t -test). A mutation is considered deleterious if its fitness is significantly lower than 1 ($P < 0.05$) and neutral if its fitness is not significantly different from 1.



Extended Data Fig. 10 | A new model explaining the widespread negative correlation between the mRNA level of a gene and its evolutionary rate measured by the nonsynonymous or amino acid substitution rate.

Compared with nonsynonymous mutations in lowly expressed genes, those in highly expressed genes tend to reduce the gene expression level and hence

tend to be deleterious. As a result, the evolutionary rate of a gene measured by the nonsynonymous or amino acid substitution rate is negatively correlated with the gene expression level. The height of a symbol represents the quantity considered.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Sequencing data generated in this study have been deposited into NCBI with the Bioproject ID of PRJNA750109. All other data are presented in the paper and associated supplementary materials. Source data are provided with this paper. Public data used include gene function annotations in the Saccharomyces Genome Database (<https://www.yeastgenome.org/>) and genomic coding sequences of *S. paradoxus*, *C. glabrata*, and *S. castellii* and genomic sequences of *S. mikatae* and *S. uvarum* from the NCBI genome assembly database (<https://www.ncbi.nlm.nih.gov/assembly/>).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	See Methods. Sample sizes were generally determined based on the goal of obtaining a good representation from the entire genome and the known levels of noise in fitness and expression quantifications.
Data exclusions	None.
Replication	3-4 replicates (all successful); see Methods for details.
Randomization	None.
Blinding	Not used, because no analysis required blinding as in clinical studies.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern		

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	BY4742 (<i>S. cerevisiae</i>)
Authentication	Recently genome sequenced
Mycoplasma contamination	No, based on genome sequencing
Commonly misidentified lines (See ICLAC register)	None.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

The yeast cells were fixed by ethanol and then treated by RNase and protease. The cell suspension was stained by SYTOX Green.

Instrument

IntelliCyt® iQue Screener PLUS

Software

FlowCore, FlowClust R packages and custom R scripts

Cell population abundance

The cells were not sorted in our method.

Gating strategy

First, the cluster analysis was used to find a major cluster based on FSC.A and SSC.A to get rid of outliers. Then, the doublets were removed according to FSC.A/FSC.H ratio.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.