LETTER

doi:10.1038/nature14244

# Selection on noise constrains variation in a eukaryotic promoter

Brian P. H. Metzger[1]*, David C. Yuan[2]*, Jonathan D. Gruber[1], Fabien Duveau[1] & Patricia J. Wittkopp[1,2]

**Genetic variation segregating within a species reflects the combined activities of mutation, selection, and genetic drift. In the absence of selection, polymorphisms are expected to be a random subset of new mutations; thus, comparing the effects of polymorphisms and new mutations provides a test for selection[1–4]. When evidence of selection exists, such comparisons can identify properties of mutations that are most likely to persist in natural populations[2]. Here we investigate how mutation and selection have shaped variation in a *cis*-regulatory sequence controlling gene expression by empirically determining the effects of polymorphisms segregating in the *TDH3* promoter among 85 strains of *Saccharomyces cerevisiae* and comparing their effects to a distribution of mutational effects defined by 236 point mutations in the same promoter. Surprisingly, we find that selection on expression noise (that is, variability in expression among genetically identical cells[5]) appears to have had a greater impact on sequence variation in the *TDH3* promoter than selection on mean expression level. This is not necessarily because variation in expression noise impacts fitness more than variation in mean expression level, but rather because of differences in the distributions of mutational effects for these two phenotypes. This study shows how systematically examining the effects of new mutations can enrich our understanding of evolutionary mechanisms. It also provides rare empirical evidence of selection acting on expression noise.**

The *TDH3* gene encodes a highly expressed enzyme involved in central glucose metabolism[6]. Deletion of this gene decreases fitness[7] and its overexpression alters phenotypes[8], suggesting that the promoter controlling its expression is subject to selection in the wild. To test this hypothesis, we sequenced a 678 base pair (bp) region containing the *TDH3* promoter ($P_{TDH3}$) as well as the 999 bp coding sequence of *TDH3* in 85 strains of *S. cerevisiae* sampled from diverse environments (Supplementary Table 1). We observed 44 polymorphisms in $P_{TDH3}$: 35 single nucleotide polymorphisms (SNPs) at 33 different sites and nine insertions or deletions (indels) ranging from 1 to 32 bp (Extended Data Fig. 1a). This frequency of polymorphic sites was significantly lower than the frequency of synonymous polymorphisms within the *TDH3* coding sequence ($P = 0.03$, Fisher's exact test) and polymorphic sites were less conserved between species than non-polymorphic sites in the promoter ($P = 5 \times 10^{-5}$, Wilcoxon rank sum test), consistent with purifying selection acting on $P_{TDH3}$. To determine whether the polymorphisms observed in $P_{TDH3}$ contribute to *cis*-regulatory variation, we compared relative *cis*-regulatory activity between each of 48 strains and a common reference strain. We found significant differences in *cis*-regulatory activity among strains (Extended Data Fig. 1b), and 97% of the heritable *cis*-regulatory variation could be explained by sequence variation within the *TDH3* promoter (see Methods). These differences in *cis*-regulation act together with differences in *trans*-regulation to produce variation in *TDH3* messenger RNA (mRNA) abundance observed among strains (Extended Data Fig. 1b).

To quantify the effect of each individual polymorphism on *cis*-regulatory activity, we used parsimony to reconstruct the evolutionary relationships among the 27 $P_{TDH3}$ haplotypes observed in the 85 strains of *S. cerevisiae* sampled. We then inferred the most likely ancestral state for these haplotypes using $P_{TDH3}$ sequences from an additional 15 strains of *S. cerevisiae* and all known species in the *Saccharomyces sensu stricto* genus (Supplementary Table 1 and Extended Data Fig. 2a). Next, we measured *cis*-regulatory activity of $P_{TDH3}$ for the inferred ancestral state, each observed haplotype, and both possible intermediates between all pairs of observed haplotypes that differed by two mutational steps. We did this by cloning each $P_{TDH3}$ haplotype upstream of the coding sequence for a yellow fluorescent protein (YFP), integrating these reporter genes ($P_{TDH3}$–*YFP*) into the *S. cerevisiae* genome, and quantifying YFP fluorescence using flow cytometry[9]. For each genotype, YFP fluorescence was measured in approximately 10,000 single cells from each of nine biological replicate populations (Fig. 1a). We used these data to estimate both mean expression level ($\mu$; Fig. 1b) and expression noise ($\sigma/\mu$; Fig. 1c) of $P_{TDH3}$–*YFP* for each promoter haplotype as readouts of *cis*-regulatory activity. We then inferred the effects of individual polymorphisms by comparing the phenotypes of ancestral and descendent haplotypes that differed by only a single sequence change.

To determine how the effects of $P_{TDH3}$ polymorphisms compare with the effects of new mutations in this *cis*-regulatory element, we estimated the distribution of mutational effects by using site-directed mutagenesis to introduce 236 of the 241 possible G:C→A:T transitions individually into $P_{TDH3}$–*YFP* alleles and assayed their effects on *cis*-regulatory activity using flow cytometry as described above. We used G:C→A:T transitions to estimate the distribution of mutational effects because they were the most common type of SNP observed both in the *TDH3* promoter (Extended Data Fig. 1a) and genome-wide among the 85 *S. cerevisiae* strains[10,11]. They were also the most frequent type of spontaneous point mutation observed in mutation accumulation lines of *S. cerevisiae*[12]. To determine whether the effects of these mutations were likely to be representative of the effects of all types of point mutation, we analysed data from previously published studies that measured the effects of single mutations on *cis*-regulatory activity[13–16]. We found no significant difference between the effects of G:C→A:T transitions and other types of point mutation on *cis*-regulatory activity in any of these data sets (Extended Data Fig. 3a–m). Consistent with this observation, we found no significant difference between the effects of G→A and C→T mutations on $P_{TDH3}$ activity (mean expression level: $P = 0.73$; expression noise: $P = 0.52$; two-tailed *t*-test; Extended Data Fig. 3n, o). We also found no significant difference between the effects of G:C→A:T and other types of polymorphism (mean expression level: $P = 0.91$; expression noise: $P = 0.90$; two-tailed *t*-test; Extended Data Fig. 3p, q).

Mutations with the largest effects on mean expression level and expression noise were located within experimentally validated transcription factor binding sites (TFBS)[17,18] (Fig. 2). All of these mutations decreased mean expression level and increased expression noise. Outside the known TFBS, 50% of the 218 mutations tested increased mean expression level and 87% increased expression noise. Despite this difference in the shape of the distributions, a negative correlation was observed between mean

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA. [2]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109, USA.
*These authors contributed equally to this work.

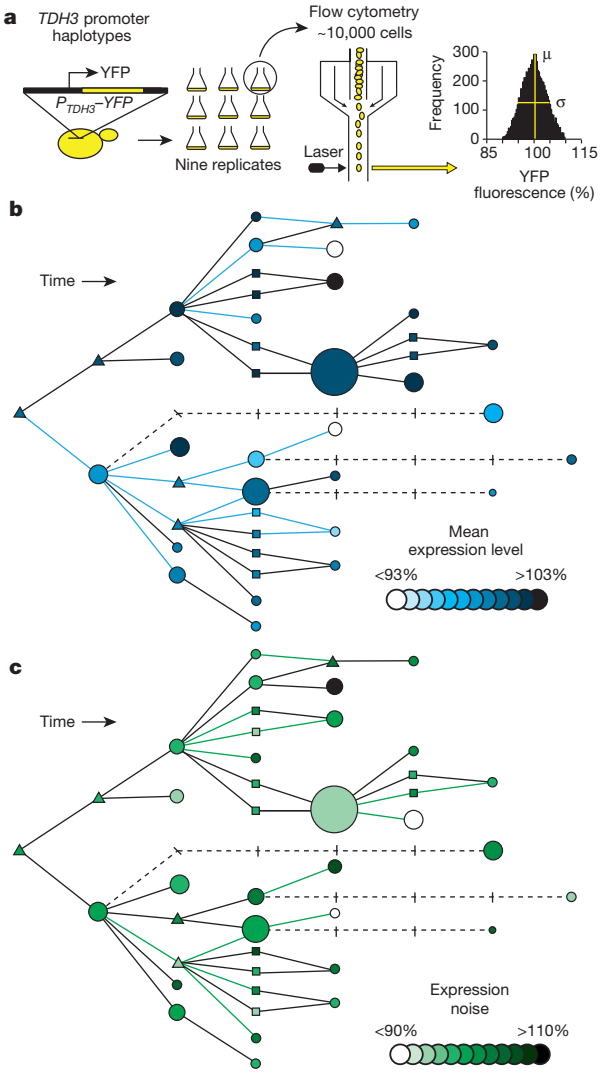



**Figure 1 | Effects of polymorphisms on $P_{TDH3}$ activity.** **a**, The *cis*-regulatory activity was quantified as YFP fluorescence in nine biological replicates for each $P_{TDH3}$–*YFP* haplotype using flow cytometry. The mean ($\mu$) and standard deviation ($\sigma$) of single-cell fluorescence phenotypes were calculated for each sample. **b**, Mean expression level of $P_{TDH3}$–*YFP* for each *TDH3* promoter haplotype is shown in the haplotype network (Extended Data Fig. 2a), with differences in mean expression level relative to the inferred common ancestor shown with different shades. Circles are haplotypes observed among the sampled strains, with the diameter of each circle proportional to frequency of that haplotype among the 85 strains. Triangles are haplotypes that were not observed among the strains sampled, but must exist, or have existed, as intermediates between observed haplotypes. Squares are possible haplotypes that might exist, or have existed, as intermediates between observed haplotypes. Dashed lines connect haplotypes by multiple mutations. On the basis of *t*-tests with a Bonferroni correction, 17 of the 45 polymorphisms present in this network caused a significant change in mean expression level (blue lines). **c**, Same as **b**, but for expression noise. Eighteen of the 45 polymorphisms present in this network caused a significant change in expression noise (green lines, *t*-test, Bonferroni corrected).

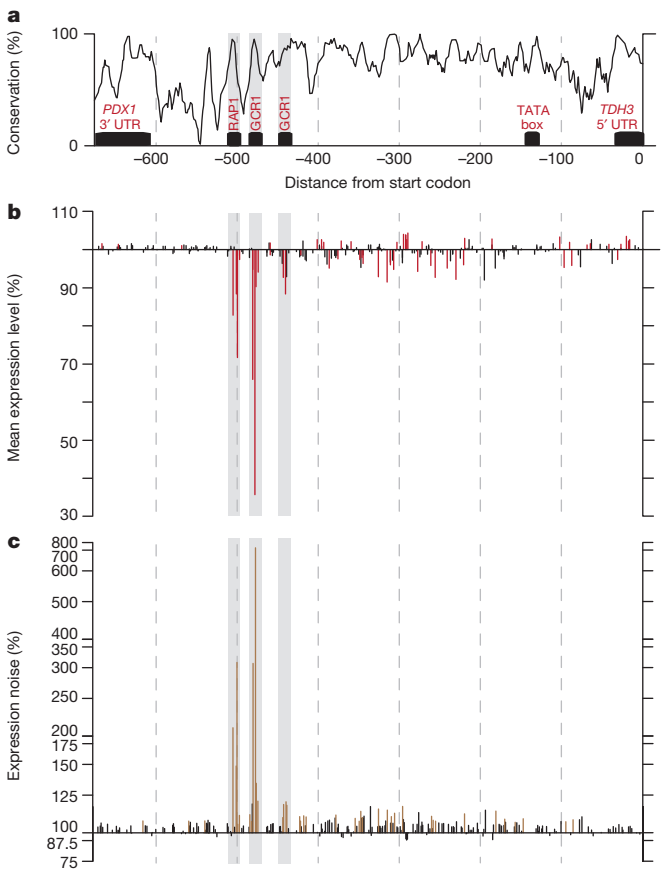



**Figure 2 | Effects of mutations on $P_{TDH3}$ activity.** **a**, The structure of the 678 bp region analysed, including the *TDH3* promoter with previously identified TFBS for RAP1 and GCR1, a TATA box, and untranslated regions (UTRs) for *TDH3* and *PDX1*, is shown. The black line indicates sequence conservation across the *sensu stricto* genus. **b**, Effects of individual mutations on mean expression level are shown in terms of the percentage change relative to the un-mutagenized reference allele, and are plotted according to the site mutated in the 678 bp region. Fifty-nine of the 236 mutations tested significantly altered mean expression levels (red lines, *t*-test, Bonferroni corrected). The shaded regions correspond to the known binding sites indicated in **a**. **c**, Same as **b**, but for expression noise. Because the effects of mutations on expression noise relative to the reference allele were much greater in magnitude than the effects of these mutations on mean expression level, they are plotted on a $\log_2$ scale. Measurements of expression noise were more variable among replicates than measurements of mean expression level, resulting in lower power to detect small changes as significant. Nonetheless, 42 of the 236 mutations tested significantly altered expression noise (brown lines, *t*-test, Bonferroni corrected).

expression level and expression noise ($R^2 = 0.85$; Extended Data Fig. 4) that was similar to previous reports for other yeast promoters[19]. The strength of this correlation was reduced to $R^2 = 0.45$ when mutations in the known TFBS were excluded.

To take the mutational process into account when testing for evidence that selection has influenced variation in the *S. cerevisiae TDH3* promoter, we compared the distributions of effects for mutations and polymorphisms both of mean expression level (Fig. 3a) and of expression noise (Fig. 3b). We did this by randomly sampling sets of variants from the mutational distribution and comparing their effects with those observed among the naturally occurring polymorphisms. We found that the effects of observed polymorphisms on mean expression level were consistent with random samples of mutations from the distribution of mutational effects (one-sided $P = 0.89$; Extended Data Fig. 5a, i), whereas the effects of observed polymorphisms on expression noise were not (one-sided $P = 0.0092$; Extended Data Fig. 5b). Specifically, polymorphisms were less likely to increase expression noise than random mutations (Extended Data Fig. 5j), suggesting that selection has preferentially retained mutations that minimize expression noise from $P_{TDH3}$ in natural populations. These results were robust to the exclusion of the large effect mutations in known TFBS from the distribution of mutational effects and the restriction of polymorphisms to G:C→A:T changes (Extended Data Fig. 5c–f, k–n), the metric used to quantify expression noise (Extended Data Fig. 6), and differences in genetic background that include a change in ploidy from haploid to diploid (Extended Data Fig. 7).

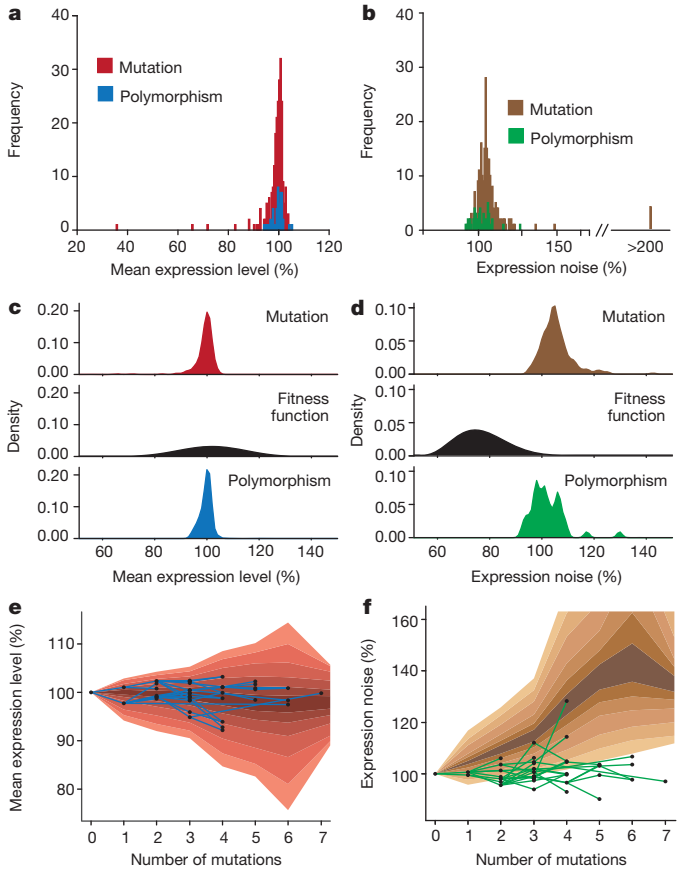



**Figure 3 | Effects of selection on $P_{TDH3}$ activity.** **a**, Summary of the effects of mutations (red) and polymorphisms (blue) on mean expression level. **b**, Summary of the effects of mutations (brown) and polymorphisms (green) on expression noise. **c**, The maximum likelihood fitness function (middle, black) relating the distribution of mutational effects (top, red) to the distribution of observed polymorphisms (bottom, blue) is shown for mean expression level. **d**, Same as **c**, but for expression noise. **e**, Changes in mean expression level observed among haplotypes over time in the inferred haplotype network (Extended Data Fig. 2a) are shown in blue. The red background represents the 95th, 90th, 80th, 70th, 60th, and 50th percentiles, from light to dark, for mean expression level resulting from 10,000 independent simulations of phenotypic trajectories in the absence of selection. **f**, Same as **e**, but for expression noise. Effects of the mutational distribution are shown in brown. Expression noise among haplotypes is shown in green.

The probability that a new mutation with a particular phenotypic effect survives within a species to be sampled as a polymorphism is related to its effect on relative fitness. The function describing relative fitness for different phenotypes can therefore be inferred by comparing the distribution of effects for new mutations to the distribution of effects for polymorphisms (Fig. 3c, d). For mean expression level, we found that the most likely fitness function (Fig. 3c) did not explain the data significantly better than a uniform fitness function representing neutral evolution ($P = 0.87$). For expression noise, we rejected a model of neutral evolution ($P = 0.00019$) and found that the most likely fitness function included higher fitness for variants that decreased gene expression noise (Fig. 3d). Repeating this analysis using alternative metrics for expression noise produced comparable results (Extended Data Fig. 6). These data suggest an evolutionary model in which purifying selection preferentially removes variants that increase expression noise, resulting in robust expression of *TDH3* among genetically identical individuals.

Consistent with this model, polymorphisms with the largest effects on expression noise (but not mean expression level) were found at the lowest frequencies within the sampled strains of *S. cerevisiae* (mean, $P = 0.43$; noise, $P = 0.0029$; permutation test; Extended Data Fig. 2b, c). However, this pattern could also result from population structure among the sampled strains. To separate the effects of selection and population structure, we used the structure of the inferred haplotype network and the distribution of mutational effects to simulate neutral trajectories for *cis*-regulatory phenotypes as they diverged from the $P_{TDH3}$ ancestral state. We then compared these trajectories with the phenotypic changes observed among naturally occurring haplotypes and their inferred intermediates both for mean expression level (Fig. 3e) and for expression noise (Fig. 3f). We found that the observed haplotypes were consistent with neutral expectations for mean expression level (one-sided $P = 0.32$; Extended Data Fig. 5g), but were not consistent with this neutral model for expression noise (one-sided $P < 0.0001$; Extended Data Fig. 5h), regardless of which metric was used to measure expression noise (Extended Data Fig. 6). We again saw that naturally occurring haplotypes showed smaller changes in noise relative to the common ancestor than would be expected from the mutational process alone, implying persistent selection for low noise in $P_{TDH3}$ activity in the wild.

Taken together, our data indicate that sequence variation in the *S. cerevisiae TDH3* promoter has been affected more by selection for low levels of noise than by selection for a particular level of *cis*-regulatory activity. This is not because the mean level of *cis*-regulatory activity is less important than noise for fitness, but because of differences in the distributions of mutational effects for these two phenotypes. Indeed, theoretical work shows that selection for low levels of noise is most likely to occur for phenotypes that are subject to purifying selection[20]. Additional evidence suggesting that selection can act on expression noise comes from genomic analyses[20–25] and from the conservation of 'shadow enhancers' that appear to maintain robust expression in multicellular organisms[26,27]. By investigating not only the survival of the fittest, but also the 'arrival of the fittest'[28,29], our work shows how phenotypic diversity produced by the mutational process itself has inherent biases that can influence the course of regulatory evolution. By taking empirical measurements of these mutational biases into account, we have identified an unexpected target of selection that impacts how a *cis*-regulatory element evolves.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 1 May 2014; accepted 19 January 2015.**
**Published online 16 March 2015.**

1. Smith, J. D., McManus, K. F. & Fraser, H. B. A novel test for selection on *cis*-regulatory elements reveals positive and negative selection acting on mammalian transcriptional enhancers. *Mol. Biol. Evol.* **30,** 2509–2518 (2013).
2. Denver, D. R. *et al.* The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nature Genet.* **37,** 544–548 (2005).
3. Stoltzfus, A. & Yampolsky, L. Y. Climbing mount probable: mutation as a cause of nonrandomness in evolution. *J. Hered.* **100,** 637–647 (2009).
4. Rice, D. P. D. & Townsend, J. P. J. A test for selection employing quantitative trait locus and mutation accumulation data. *Genetics* **190,** 1533–1545 (2012).
5. Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304,** 1811–1814 (2004).
6. McAlister, L. & Holland, M. J. Differential expression of the three yeast glyceraldehyde-3-phosphate dehydrogenase genes. *J. Biol. Chem.* **260,** 15019–15027 (1985).
7. Pierce, S. E., Davis, R. W., Nislow, C. & Giaever, G. Genome-wide analysis of barcoded *Saccharomyces cerevisiae* gene-deletion mutants in pooled cultures. *Nature Protocols* **2,** 2958–2974 (2007).
8. Ringel, A. E. *et al.* Yeast Tdh3 (glyceraldehyde 3-phosphate dehydrogenase) is a Sir2-interacting factor that regulates transcriptional silencing and rDNA recombination. *PLoS Genet.* **9,** e1003871 (2013).
9. Gruber, J. D., Vogel, K., Kalay, G. & Wittkopp, P. J. Contrasting properties of gene-specific regulatory, coding, and copy number mutations in *Saccharomyces cerevisiae*: frequency, effects and dominance. *PLoS Genet.* **8,** e1002497 (2012).
10. Liti, G. *et al.* Population genomics of domestic and wild yeasts. *Nature* **458,** 337–341 (2009).
11. Schacherer, J., Shapiro, J. A., Ruderfer, D. M. & Kruglyak, L. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* **458,** 342–345 (2009).
12. Lynch, M. *et al.* A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl Acad. Sci. USA* **105,** 9272–9277 (2008).
13. Patwardhan, R. P. *et al.* High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nature Biotechnol.* **27,** 1173–1175 (2009).

14. Melnikov, A. *et al.* Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nature Biotechnol.* **30,** 271–279 (2012).
15. Patwardhan, R. P. *et al.* Massively parallel functional dissection of mammalian enhancers *in vivo*. *Nature Biotechnol.* **30,** 265–270 (2012).
16. Kwasnieski, J. & Mogno, I. Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc. Natl Acad. Sci. USA* **109,** 19498–19503 (2012).
17. Yagi, S., Yagi, K., Fukuoka, J. & Suzuki, M. The UAS of the yeast GAPDH promoter consists of multiple general functional elements including RAP1 and GRF2 binding sites. *J. Vet. Med. Sci.* **56,** 235–244 (1994).
18. Baker, H. V. *et al.* Characterization of the DNA-binding activity of GCR1: *in vivo* evidence for two GCR1-binding sites in the upstream activating sequence of *TPI* of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12,** 2690–2700 (1992).
19. Hornung, G. *et al.* Noise-mean relationship in mutated promoters. *Genome Res.* **22,** 2409–2417 (2012).
20. Lehner, B. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.* **4,** 1–6 (2008).
21. Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J. & Eisen, M. B. Noise minimization in eukaryotic gene expression. *PLos Biol.* **2,** 834–838 (2004).
22. Wang, Z. & Zhang, J. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc. Natl Acad. Sci. USA* **108,** E67–E76 (2011).
23. Newman, J. R. S. *et al.* Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441,** 840–846 (2006).
24. Batada, N. & Hurst, L. Evolution of chromosome organization driven by selection for reduced gene expression noise. *Nature Genet.* **39,** 945–949 (2007).
25. Zhang, Z., Qian, W. & Zhang, J. Positive selection for elevated gene expression noise in yeast. *Mol. Syst. Biol.* **5,** 1–12 (2009).
26. Frankel, N. *et al.* Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466,** 1–5 (2010).
27. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20,** 1562–1567 (2010).
28. Fontana, W. & Buss, L. "The arrival of the fittest": toward a theory of biological organization. *Bull. Math. Biol.* **56,** 1–64 (1994).
29. De Vries, H. *Species and Varieties, Their Origin by Mutation* 825–826 (Open Court, 1905).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank C. Maclean, J. Zhang, and C. Hittinger for strains, the University of Michigan Center for Chemical Genomics for technical assistance with flow cytometry, and J. Coolon, R. Lusk, K. Stevenson, A. Hodgins-Davis, J. Lachowiec, C. Maclean, J. Yang, C. Landry, J. Townsend, and D. Petrov for comments on the manuscript. Funding for this work was provided to P.J.W. by the March of Dimes (5-FY07-181), Alfred P. Sloan Research Foundation, National Science Foundation (MCB-1021398), National Institutes of Health (1 R01 GM108826) and the University of Michigan. Additional support was provided by the University of Michigan Rackham Graduate School, Ecology and Evolutionary Biology Department, and the National Institutes of Health Genome Sciences training grant (T32 HG000040) to B.P.H.M.; by a National Institutes of Health Genetics training grant (T32 GM007544) to D.C.Y.; by a National Institutes of Health National Research Service Award (NRSA) postdoctoral fellowship (1 F32 GM083513-0) to J.D.G.; and by a European Molecular Biology Organization postdoctoral fellowship (EMBO ALTF 1114-2012) to F.D.

**Author Contributions** D.C.Y., P.J.W., and J.G. designed the mutational spectrum project. D.C.Y. created all $P_{TDH3}$–*YFP* mutant strains. D.C.Y., B.P.H.M., and F.D. collected flow cytometry data. B.P.H.M. and P.J.W. designed the natural variation project. D.C.Y. created all strains with natural haplotypes and B.P.H.M. performed all other experiments. B.P.H.M. analysed all data. B.P.H.M., D.C.Y., and P.J.W. wrote the manuscript.

**Author Information** Flow cytometry data have been deposited in the FlowRepository under Repository ID FR-FCM-ZZBN. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.J.W. (wittkopp@umich.edu).

## METHODS

**Characterizing variation segregating in the *TDH3* promoter.** Variation in the *TDH3* gene was determined for 85 natural isolates of *S. cerevisiae*[10,11] (Supplementary Table 1). Sequences were obtained from each strain by PCR and Sanger sequencing using DNA extracted from diploid cells. Strains heterozygous for the *TDH3* promoter were grown on GNA plates for 12 h (5% dextrose, 3% Difco nutrient broth, 1% Oxoid yeast extract, 2% agar) and sporulated on potassium acetate plates (1% potassium acetate, 0.1% Oxoid yeast extract, 0.05% dextrose, 2% agar). Individual spores were isolated by tetrad dissection and haploid derivatives were sequenced to determine empirically the phase of the two *TDH3* promoter haplotypes. All reagents for growth of yeast cultures were purchased from Fisher unless otherwise noted. In all, the 678 bp promoter contained SNPs at 33 sites and the 238 synonymous sites contained 22 SNPs. Five non-synonymous changes were also observed among these 85 strains.

**Inferring the ancestral sequence and constructing the haplotype network for $P_{TDH3}$.** Promoter haplotypes (Supplementary Table 1 and Extended Data Fig. 2a) were initially aligned using Pro-Coffee[30], followed by re-alignment with PRANK[31] and manual adjustment around repetitive elements and indels (Supplementary File 1). The *TDH3* promoter sequences from all *Saccharomyces sensu stricto* species[10,32–34], as well as an additional 15 strains of *S. cerevisiae* known to be an outgroup to the 85 focal strains[35], were also determined by Sanger sequencing. These sequences were used to infer the ancestral state of the *TDH3* promoter for the 85 strains with both parsimony and maximum likelihood methods implemented in MEGA 6 (ref. 36); both methods gave identical results. TCS 2.1 (ref. 37) was used to build a haplotype network for the *TDH3* promoter, with changes polarized on the basis of the inferred ancestral state (Extended Data Fig. 2a). One haplotype (HH in Supplementary Table 1) could not be confidently placed within the network and was excluded from our analysis. Sequence conservation for individual sites was determined using sequences from all seven *Saccharomyces sensu stricto* species using ConSurf[38] and the phylogeny from a prior study[39]. To reduce heterogeneity in plotting, conservation was averaged over a 20 bp sliding window.

**Measuring variation in *TDH3* mRNA levels and *cis*-regulatory activity.** *TDH3* mRNA levels and *cis*-regulatory activity were measured using pyrosequencing, with relative allelic expression in $F_1$ hybrids providing a readout of relative *cis*-regulatory activity[40]. This technique requires one or more sequence differences to compare relative genomic DNA (gDNA) or complementary DNA (cDNA) abundance between two strains or two alleles within the same strain[41]. We therefore constructed reference strains of both mating types that carried a copy of the *TDH3* gene with a single, synonymous mutation (T243G). These genotypes were constructed by inserting the *URA3* gene into the native *TDH3* coding region in strains BY4741 and BY4742 and then replacing *URA3* with the modified *TDH3* coding sequence using the lithium acetate method and selection on 5-FOA[9,42]. To do this, 80 bp oligonucleotides, containing a synonymous mutation and homology to each side of the target site, were transformed into these strains. Successful transformants (strains YPW342 and YPW339, respectively) were confirmed by Sanger sequencing. Resistance markers for hygromycin B (*hphMX6*) and G418 (*kanMX4*) were then inserted into the *HO* locus of these strains (producing YPW360 and YPW361, respectively) and used to construct a diploid reference strain (YPW362). A *kanMX4* resistance marker was also successfully inserted into the *HO* locus of 63 of the 85 natural strains[10,11].

To construct hybrids suitable for measuring *cis*-regulatory activity of natural isolates relative to a reference strain, haploid cells from each of the 63 natural isolates with a *kanMX4* resistance marker (mating type **a**) were mixed with an equal number of haploid cells from the reference strain YPW360 (mating type α) on YPD plates (2% dextrose, 1% Oxoid yeast extract, 2% Oxoid peptone, 2% agar). After 24 h, cultures were streaked on YPD plates to obtain single colonies and then patched to YPD plates containing G418 and Hygromycin B to select for diploids. Four replicates of each hybrid were grown in 500 µl of YPD liquid media for 20 h at 30 °C in 2 ml 96-well plates with 3 mm glass beads, shaking at 250 rpm. Cultures were diluted to an attenuance, $D_{600\,nm}$, of 0.1 and then grown for an additional 4 h. Plates were centrifuged, and the YPD liquid was removed. Cultures were then placed in a dry ice/ethanol bath until frozen and stored at −80 °C. To prepare samples for measuring total *TDH3* mRNA abundance in each natural isolate relative to a common reference strain, diploids for each of the 63 natural isolates were mixed with a similar number of diploid cells from strain YPW362 on the basis of OD600 readings after the initial growth in YPD liquid. These co-cultures were incubated and processed as described above.

For each hybrid and co-culture sample, gDNA and RNA were sequentially extracted from a single lysate using a modified protocol of Promega's SV Total RNA Isolation System. After thawing cultures on ice for about 30 min, 175 µl of SV RNA lysis buffer (with β-mercaptoethanol), 350 µl of double-distilled water, and 50 µl of 400 micron RNase free beads were added to each sample. Plates were vortexed until cell pellets were completely resuspended. The plates were then centrifuged and 175 µl of supernatant was mixed with 25 µl of RNase-free 95% ethanol and loaded onto a binding plate. To extract RNA, 100 µl of RNase-free 95% ethanol was added to the flow through and loaded onto a second binding plate. These plates were then washed twice with 500 µl of SV RNA wash solution and allowed to dry. To extract DNA, the first binding plate was washed twice with 700 µl of cold 70% ethanol and allowed to dry. For both binding plates, 100 µl of double-distilled water was added to each well, the plate was incubated at 25 °C for 7.5 min, and the elution was collected. RNA from each sample was converted to cDNA by mixing 5 µl of extracted RNA with 2 µl RNase-free water, 1 µl DNase buffer, 1 µl RNasin Plus, and 1 µl DNase 1 and incubating at 37 °C for 1 h followed by 65 °C for 15 min. Three microlitres of oligo dT ($T_{19}VN$) was added and cooled to 37 °C over 35 min. Four microlitres of First Strand Buffer, 2 µl dNTPs, 0.5 µl RNasin Plus, and 0.5 µl of SuperScript II were added and incubated for 1 h. Thirty microlitres of double-distilled water was then added to each sample.

Pyrosequencing was performed as described previously[41] using a PSQ 96 pyrosequencing machine and Qiagen pyroMark Gold Q96 reagents for gDNA and cDNA samples both for hybrids and for co-cultured diploids. One microlitre of cDNA or gDNA was used in each PCR reaction, with primers shown in Supplementary Table 2. A single PCR and pyrosequencing reaction was performed for each gDNA and cDNA sample from each of the four biological replicate hybrid and co-culture samples for each natural haplotype, for a total of eight pyrosequencing reactions using cDNA and eight pyrosequencing reactions using gDNA for each of the 48 strains (Supplementary Table 3).

In gDNA samples from hybrids, the two *TDH3* alleles are expected to be equally abundant; however, differences in PCR amplification of the two alleles (or aneuploidies altering copy number of *TDH3*) can cause unequal representation in the pyrosequencing data. Because such deviations cause estimates of relative allelic expression for these samples to be less reliable, the 15% of samples with gDNA ratios that deviated by more than 15% from the expected 50:50 ratio were excluded. Relative abundance of the two *TDH3* alleles is expected to be more variable in the co-cultured samples because of unequal representation from differences in concentration of the two genotypes before mixing and/or after growth. Samples from co-cultured diploids with gDNA ratios in the upper or lower 10th percentiles were also excluded from analysis. These quality control filters left 48 strains with at least two replicates in both the hybrid and co-cultured samples.

For each sample, relative allelic abundance in the cDNA sample was divided by relative allelic abundance for the corresponding gDNA sample to correct for remaining biases[41]. These ratios ($Y_{ijk}$) from strain *i*, plate *j*, and replicate *k* were fitted to the following linear model, including strain (ranging from 1 to 48) and plate (ranging from 1 to 3) as fixed effects as well as the cell density of the sample before and after growth from which the RNA and DNA were extracted (measured by $D_{600\,nm}$) as a covariate: $Y_{ijk} = \mu + \text{strain} + \text{plate} + \text{density.0} + \text{density.1} + \varepsilon$. An analysis of variance (ANOVA) found that strain, plate, and initial density were statistically significant for hybrids (strain: $P = 1.38 \times 10^{-20}$; plate: $P = 1.01 \times 10^{-10}$; density.0: $P = 5.01 \times 10^{-3}$; density.1: $P = 0.740$), and strain and plate were statistically significant for co-cultured diploids (strain: $P = 8.16 \times 10^{-20}$; plate: $P = 2.65 \times 10^{-3}$; density.0: $P = 0.734$; density.1: $P = 0.833$). Expression values for each sample were adjusted to remove the effects of plate and initial cell density. Differences in allelic abundance caused by the synonymous change introduced for pyrosequencing were estimated by analysing a hybrid between BY4741 and YPW360 and a co-culture of BY4741 and YPW362. The effects of this change were then subtracted from the $\log_2$-transformed expression ratio for all samples. Strains with significant *cis*-regulatory divergence from the reference were identified using *t*-tests. R code used for these analyses is provided in Supplementary File 2.

To determine the amount of variation in *TDH3 cis*-regulatory activity explained by strain identity and the *TDH3* promoter haplotype, we fitted the normalized expression values to linear models containing fixed effects of either strain identity or promoter haplotype alone. Variance among strains explained by strain identity was assumed to reflect heritable variation, with residual variance assumed to result from technical noise. Because multiple strains contained the same *TDH3* promoter haplotype, we were able to determine the proportion of this heritable variance explained by polymorphisms in the *TDH3* promoter region tested. Seventy-five per cent of all *cis*-regulatory variation and 97% of heritable *cis*-regulatory variation were explained by the *TDH3* promoter haplotype. To estimate the error associated with these estimates of variance explained, we analysed 100,000 bootstrap replicates of the data with the same linear models.

**Constructing strains with mutations and polymorphisms in $P_{TDH3}$.** To assay *cis*-regulatory activity of the *TDH3* promoter efficiently, we used a $P_{TDH3}$–*YFP* reporter gene integrated near a pseudogene on chromosome 1 of strain BY4724 at position 199270 (ref. 9). This $P_{TDH3}$–*YFP* transgene contained a 678 bp sequence including the *TDH3* promoter that was fused to the coding sequence for YFP and the *CYC1* (cytochrome *c* isoform 1) terminator. The 678 bp sequence extended 5′ from the start codon of *TDH3* into the 3′ UTR of the neighbouring gene (*PDX1*),

including the 5′ UTR of *TDH3*. To facilitate replacing this reference haplotype with other $P_{TDH3}$ haplotypes, we used homologous replacement to create a derivative of this starting strain in which the $P_{TDH3}$ sequence as well as the start codon of YFP was replaced with the *URA3* gene (*URA3–YFP*; strain YPW44).

To assess *cis*-regulatory activity of naturally occurring $P_{TDH3}$ haplotypes, we amplified the *TDH3* promoters from the 85 natural isolates using PCR and transformed these PCR products into the *URA3–YFP* intermediate. Unobserved intermediate haplotypes between all pairs of haplotypes that differed at exactly two sites were constructed by PCR-mediated site-directed mutagenesis of one of the two haplotypes in each pair and transformed into the *URA3–YFP* strain. The 236 mutant $P_{TDH3}$ alleles analysed, each containing a single G:C→A:T transition, were also constructed using PCR-mediated site-directed mutagenesis, but starting with the reference $P_{TDH3}$ haplotype. Each of these sequences was also transformed into the same *URA3–YFP* strain. All PCR primers used for amplification and site-directed mutagenesis are shown in Supplementary Table 2. In all cases, (1) transformations were performed using the lithium acetate method[42]; (2) transformants were selected on 5-FOA plates, streaked for single colonies, and confirmed to not be petite (missing mitochondrial DNA) by replica plating onto YPG plates (3% (v/v) glycerol, 2% Oxoid yeast extract, 2% Oxoid peptone, 2% agar); and (3) Sanger sequencing was used to determine the sequence of potential transformants.

**Quantifying fluorescence of $P_{TDH3}$–*YFP*, a proxy for *cis*-regulatory activity of $P_{TDH3}$.** Prior work shows that fluorescence of reporter proteins such as YFP provide a reliable readout of *cis*-regulatory activity[9,43]. Before quantifying fluorescence, all strains were revived from glycerol stocks onto YPG at the same time to control for age related effects on expression. Strains were inoculated from YPG solid media into 500 µl of YPD liquid media and grown for 20 h at 30 °C in 2 ml 96-well plates with 3 mm glass beads, shaking at 250 rpm. Immediately before flow cytometry, 20 µl of the overnight culture was transferred into 500 µl of SC-R (dextrose) media[9]. Flow cytometry data were collected on an Accuri C6 using an Intellicyt Hypercyt Autosampler. Flow rate was 14 µl $min^{-1}$ and core size was 10 µm. A blue laser ($\lambda = 488$ nm) was used for excitation of YFP. Data were collected from FL1 using a 533/30 nm filter. Each culture was sampled for 2–3 s, resulting in approximately 20,000 recorded events.

Samples were processed using the flowClust[44] and flowCore[45] packages within R (version 3.0.2) and custom R scripts[46] (http://www.r-project.org/) (Supplementary File 3). Raw data (Extended Data Fig. 8a) were $\log_{10}$ transformed and artefacts were removed by excluding events with extreme FSC.H, FSC.A, SSC.H, SSC.A, and width values (Extended Data Fig. 8b). Samples were clustered on the basis of FSC.A and width to remove non-viable cells and cellular debris, then clustered on FSC.H and FSC.A to remove doublets (Extended Data Fig. 8c). Finally, samples were clustered on FL1.A and FSC.A to obtain homogeneous populations of cells in the same stage of the cell cycle (Extended Data Fig. 8d). At each filtering step, data were divided into exactly two clusters. Samples containing fewer than 1,000 events after processing were discarded. For each sample, YFP expression was calculated as the median $\log_{10}(FL1.A)^2/\log_{10}(FSC.A)^3$. This corrected YFP expression levels for the correlation between fluorescence and cell size (measured by FSC.A) (Extended Data Fig. 8e). Expression noise for each sample was calculated as $\sigma/\mu$. The following alternative metrics for expression noise were also calculated and used for analysis: $\sigma$, $\sigma^2/\mu^2$, $\sigma^2/\mu$, and residuals from a regression of $\sigma$ on $\mu$.

For each genotype, nine independent replicate cultures were analysed, with three biological replicates included on each of three different days. Power analyses indicated that six replicates were sufficient to detect differences in mean expression of 2% at $\alpha = 0.05$ and power greater than 90%. To control for variation in growth conditions, all plates contained 20 replicates of the wild-type reference strain, with at least one control sample in each row and column of the plate. For both mean expression and the standard deviation of expression, the control samples were fitted to a linear model that included final cell number and average cell width as well as the day, replicate, array, read order, growth position in the incubator, array depth in incubator, measurement block, row, and column of the sample. Stepwise Akaike information criterion was performed on this model to identify the most informative combination of variables to keep in the model. Plate (which incorporated effects of day, replicate, and array) and block were significant from this model. The effects of these factors were removed from measures of YFP (Extended Data Fig. 8f–y) before the final analysis. A non-fluorescent strain containing no *TDH3* promoter was used to estimate autofluorescence and this value was subtracted from all YFP expression values (Supplementary File 4 and Supplementary Table 4).

The effect of an individual polymorphism on mean expression level and expression noise was measured as the difference in phenotype between the descendant and ancestral haplotypes that varied only for that polymorphism. The effect of an individual mutation on mean expression level and expression noise was measured as the difference in phenotype between the reference strain and the strain carrying that mutation. Statistical significance of effects for individual polymorphisms and mutations was assessed using two-sided *t*-tests.

Although we frequently switched to fresh clones from glycerol stocks of the *URA3–YFP* strain during construction of the collection of 381 $P_{TDH3}$–*YFP* strains analysed in this study, we checked for the presence of relevant second-site mutations that might have arisen spontaneously by independently reintroducing the $P_{TDH3}$ reference allele three times. No difference in YFP fluorescence was observed among these replicate strains for either mean expression level or expression noise (mean $P = 0.16$, noise $P = 0.069$, $n = 1{,}483$, ANOVA).

The reference haplotype used to determine the effect of new mutations differs from the most closely related natural haplotype (haplotype A) by a single base pair. To determine the impact of this single nucleotide difference on the distribution of mutational effects for mean expression level and expression noise, we introduced 28 of the G:C→A:T mutations into haplotype A and constructed $P_{TDH3}$–*YFP* strains that carried these alleles. The 28 mutations chosen for testing showed a range of effects on both mean expression level and expression noise. We found that this single base difference significantly decreased mean expression level by 3.7% ($P = 8.1 \times 10^{-56}$, ANOVA) and significantly increased expression noise by 6.8% ($P = 1.61 \times 10^{-4}$, ANOVA), but these effects were largely consistent across genetic backgrounds, indicating little and/or weak epistasis (Extended Data Fig. 9a, b). Indeed, we found that the distributions of mutational effects estimated by these 28 mutations on haplotype A and the 236 mutations on the reference haplotype were similar for both mean expression level and expression noise (Extended Data Fig. 9c, d).

The reference background also contained 6 bp at the 5′ end of the $P_{TDH3}$ region derived from the 3′ UTR of *PDX1* that was not included in the $P_{TDH3}$–*YFP* constructs containing natural $P_{TDH3}$ haplotypes. To determine whether this sequence was likely to have affected our measurements of polymorphism effects, we tested for a significant change in YFP fluorescence when these 6 bp were added to the $P_{TDH3}$–*YFP* alleles carrying the natural haplotypes A, D, and VV. We found no significant difference between genotypes with and without this 6 bp sequence (mean $P = 0.88$, noise $P = 0.25$, ANOVA).

To determine the sensitivity of our conclusions to the specific genetic background used to assay *cis*-regulatory activity, we created hybrids between one of the natural *S. cerevisiae* isolates (YPS1000) and (1) 111 strains with mutations in $P_{TDH3}$–*YFP*, (2) the strain carrying the reference $P_{TDH3}$–*YFP* allele, (3) 39 strains with naturally occurring *TDH3* promoter haplotypes driving YFP expression, and (4) a strain without the *TDH3* promoter in the $P_{TDH3}$–*YFP* construct and thus no YFP expression. YPS1000 was isolated from an oak tree and is substantially diverged from strain BY4724 (>53,000 SNPs, 0.44% (refs 10, 11)). We crossed all 152 of the strains described above (mating type **a**) to an isolate of YPS1000 that contained a KanMX4 drug resistance marker at the HO locus (mating type α). Hybrids were created by mixing equal cell numbers in liquid YPD and growing at 30 °C for 48 h without shaking. Cultures were diluted and plated on YPG + G418 to select for hybrids and prevent petite cells from growing. Colonies were grown for 48 h and then screened by fluorescence microscopy for YFP expression. Fluorescent colonies were streaked for single colonies and then a single colony was randomly chosen from each plate, transferred to a new plate, and confirmed to be diploid using a PCR reaction that genotyped the mating type locus. Four replicates of each strain were arrayed as in the original experiment with 20 controls per 96-well plate. Samples were grown for 20 h in 500 µl of YPD liquid with shaking at 30 °C, then analysed using the same flow cytometer machine and conditions described above. Samples were processed using the same analysis scripts described above, and mean expression level and expression noise were calculated. Eight of the 111 genotypes carrying reporter genes with mutations as well as four of the 39 genotypes carrying reporter genes with polymorphisms showed phenotypes suggesting that they were aneuploidies. This rate was consistent with our previous observations of spontaneous aneuploidies produced by BY4742 (ref. 9). One additional strain (containing a mutation in the *TDH3* promoter) was also excluded for having highly inconsistent measurements among replicate populations. The R script used for this analysis is provided in Supplementary File 5 and the data are provided in Supplementary Table 5.

**Tests for evidence of natural selection.** In the absence of selection, the effects of polymorphisms are expected to be consistent with the effects of a random sample of new mutations. Because our data were non-normally distributed, we used non-parametric tests based on sampling to assess significance. To estimate the probability of occurrence for a mutation with a particular effect ($x$), we used a Gaussian kernel with a bandwidth of 0.01 to fit density curves to the distributions of mutational effects observed both for mean expression level and for expression noise. We calculated the density for mean expression level values ranging from 0% to 200%, and for expression noise values ranging from 0% to 800%, ranges that extended beyond all observed effects. We set the minimum density for any effect size to 1/(number of mutations included in the mutational distribution). We expected this minimum to overestimate the true probability of most unobserved effect sizes, making this a conservative baseline for testing whether the effects of observed polymorphisms were a biased subset of all possible mutations. These density curves

were then converted into probability distributions by setting the total density equal to 1 (Extended Data Fig. 10a, b).

To calculate the log-likelihood of a set of $n$ genetic variants with effects $x_1, x_2, \ldots, x_n$, we used these probability distributions to estimate the log-likelihood of a mutation with that effect, $p(x)$, and summed probabilities for all genetic variants. That is, the log-likelihood of a set of particular effects was calculated as $\sum_{i=1}^{n} \log(p(x_i))$. The log-likelihood calculated for the 45 observed polymorphisms was compared with the log-likelihoods of 100,000 samples of 45 mutations drawn randomly from the corresponding mutational distribution with replacement. To test the hypothesis that the effects of observed polymorphisms were unlikely to result by chance from the mutational process alone, one-sided $P$ values were calculated as the proportion of random samples with log-likelihoods less than the log-likelihood value calculated for the observed polymorphisms. To determine the effects of mutations in the known TFBS on this test for selection, we excluded the effects of the mutations in the known TFBS from the distribution of mutational effects, recalculated the density curves and probability distributions, then recalculated the log-likelihoods and $P$ values.

Fitness functions relate the effect of a new mutation to its likelihood of survival within a population. We determined the most likely fitness function for mean expression level and expression noise by using a hill-climbing algorithm to identify the $\alpha$ and $\beta$ parameters of a beta distribution that maximized the likelihood of the observed polymorphism data when multiplied by the distribution of mutational effects. The beta function was started with parameters consistent with neutral evolution ($\alpha = 0, \beta = 0$) and new parameters were sampled randomly from a uniform distribution. The likelihood of the observed data was then calculated under the combined distribution of mutational effects and the new beta distribution. If the likelihood increased, the new parameters were kept; if not, they were discarded. This process was repeated until we observed 1,000 successive rejections. After each rejection, the width of the uniform distribution was increased to sample values farther away from the current parameters. A likelihood ratio test (two degrees of freedom) comparing the fitness function described by the maximum likelihood parameters for the beta distribution with a fitness function consistent with neutrality ($\alpha = 0, \beta = 0$) was used to test for statistically significant evidence of selection.

If the effects of polymorphisms are determined solely by mutation, phenotypes should drift over evolutionary time in a manner dictated by the mutational process. We modelled such a neutral scenario by starting with the phenotype of the inferred common ancestor and adding to it effects randomly drawn from the mutational distribution (sampled with replacement) for each new polymorphism observed in the haplotype network, maintaining the observed relationships among haplotypes. This process was repeated 10,000 times to generate a range of potential outcomes consistent with neutral evolution of $P_{TDH3}$ activity. We then compared the observed polymorphism data with the results of these neutral simulations to test for a statistically significant deviation from neutrality that would indicate selection. A more detailed description of this method follows.

Let $x$ be the number of new polymorphisms added to the population to convert an observed haplotype into the most closely related descendent haplotype in each lineage that exists or must have existed in wild populations of *S. cerevisiae*. In the haplotype network for $P_{TDH3}$, $x$ ranges from 0 to 5 (Extended Data Fig. 2a). Pairs of haplotypes separated by 0 new polymorphisms result from recombination between existing haplotypes (for example, haplotype RR, which is a recombinant of haplotypes W and FF).

The probability of a polymorphism with any particular effect being added to the population was assumed, in the absence of selection, to be equal to the probability of a new mutation with that effect. The log-likelihood of a single mutation ($x = 1$) with a particular effect was calculated using the probability distributions fitted to density curves based on the observed mutational distributions described above. To generate equivalent probability distributions for sets of $x = 2$, 3, 4, or 5 new mutations, we randomly drew $x$ mutations from the observed distribution of single mutational effects with replacement, calculated the combined effect of these mutations, and repeated this process 10,000 times. We then fitted a density curve to these 10,000 combined effect values for each value of $x$, set the total density to 1 to convert this into a probability distribution, and used these curves (Extended Data Fig. 10c, d) to calculate the log-likelihood of a particular set of $x$ new polymorphisms with a given combined effect in the absence of selection. A likelihood of 1 was assigned to pairs of haplotypes separated only by recombination ($x = 0$), because the new genetic variant incorporated into the descendant haplotype was already known to have arisen in the population.

To calculate an overall log-likelihood for the observed set of polymorphisms, we summed the log-likelihood values for phenotypic differences observed between each pair of most closely related haplotypes seen among the natural isolates. To determine whether this overall log-likelihood for the observed polymorphisms was consistent with neutrality, we used the structure of the haplotype network to simulate 10,000 alternative sets of haplotype effects assuming that the effect of each new polymorphism was drawn randomly from the distribution of mutational effects. We calculated the log-likelihood for each node, in each set of haplotype effects, as $\log\left[\prod_{x=1}^{5}(n_x! \times \prod_{i=1}^{n_x} p(x_i))\right]$, where $x$ is the number of mutational steps, $n_x$ is the number of immediately descendent haplotypes that are $x$ mutational steps away from the focal node that exist or must have existed in *S. cerevisiae* (Extended Data Fig. 2a), and $p(x_i)$ is the likelihood of the $i$th mutation drawn from the probability distribution based on sets of $x$ mutations. The $n_x!$ factor accounts for all possible ways that $x$ mutations (or polymorphisms) added to the population at any given step could have been arranged among the set of descendent haplotypes observed.

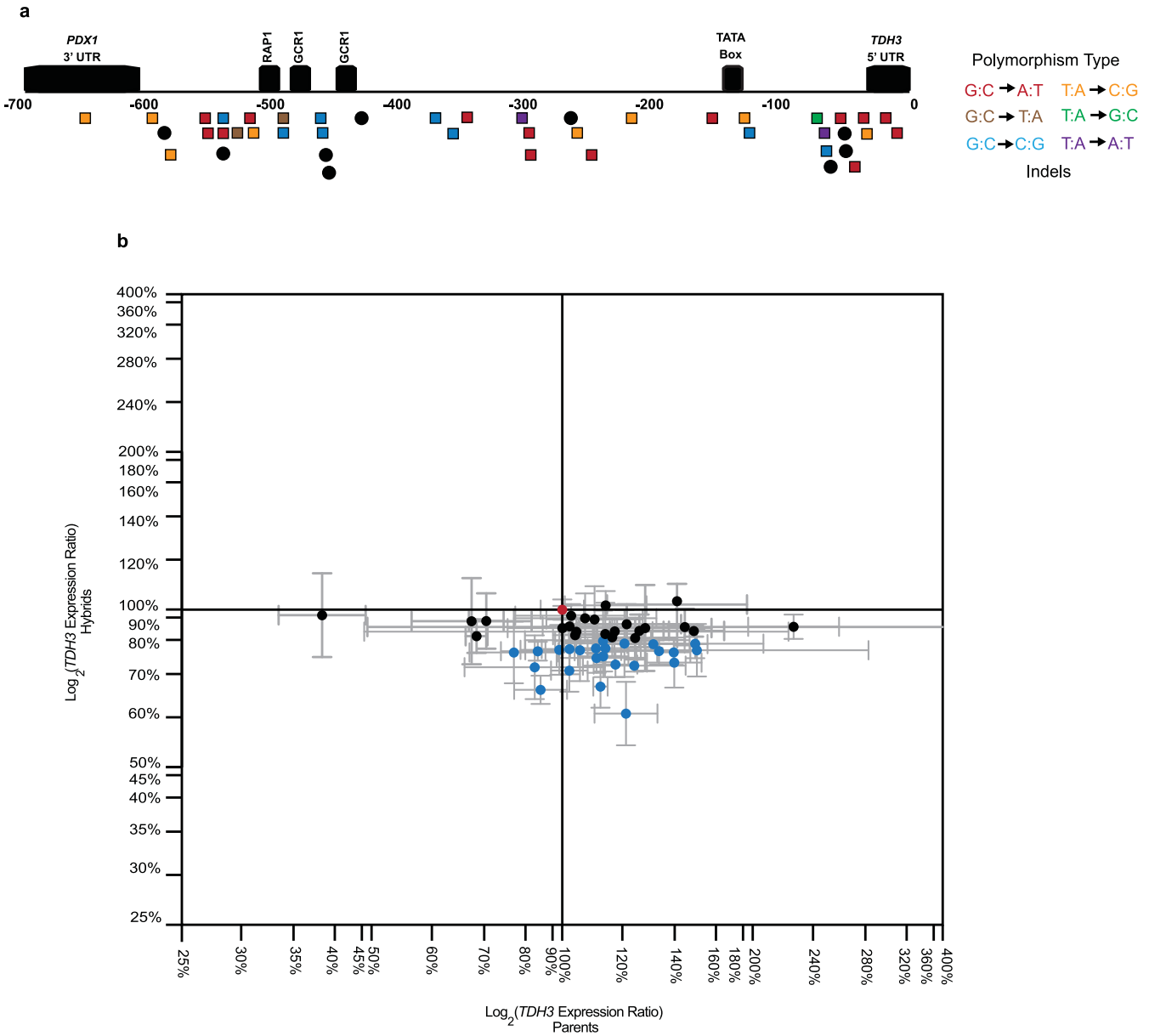
To illustrate how this works for one particularly complex node in the network, consider haplotype H and its six immediately descendent haplotypes, L, I, VV, D, S, and N (Extended Data Fig. 2a). Five of these descendent haplotypes (all except L) are all one mutational step away from H. To simulate the neutral evolution of these five haplotypes, we drew five mutational effects randomly from the probability distribution for single mutations ($x = 1$) with replacement, then determined the likelihood of each of these mutational effects based on the probability distribution for $x = 1$. These likelihood values were multiplied together to calculate the combined probability of that particular set of five mutational effects occurring. This product was then multiplied by the 5! ways in which these mutations could have been arranged among the five descendent haplotypes. We also took into account that haplotype H has one additional descendent haplotype that is five mutational steps away from H (with none of the intermediate haplotypes known) by drawing a single value randomly from the distribution of mutational effects derived from random sets of five mutations ($x = 5$); we calculated its likelihood using the probability distribution for $x = 5$; and we multiplied it by the 1! way in which this set of five mutational effects could have been added to haplotype H to produce haplotype L.

The log-likelihoods for all nodes in the haplotype network were then summed to compute the log-likelihood of each set of haplotypes. To determine whether the *cis*-regulatory phenotypes observed among the natural isolates were consistent with neutral evolution, we compared the log-likelihood calculated for the observed polymorphisms with the log-likelihoods calculated for the 10,000 data sets simulated assuming neutrality. A one-sided $P$ value was calculated as the proportion of simulated neutral data sets that had a log-likelihood value less than the log-likelihood for the observed polymorphisms (Extended Data Figs 5g, h and 6q).

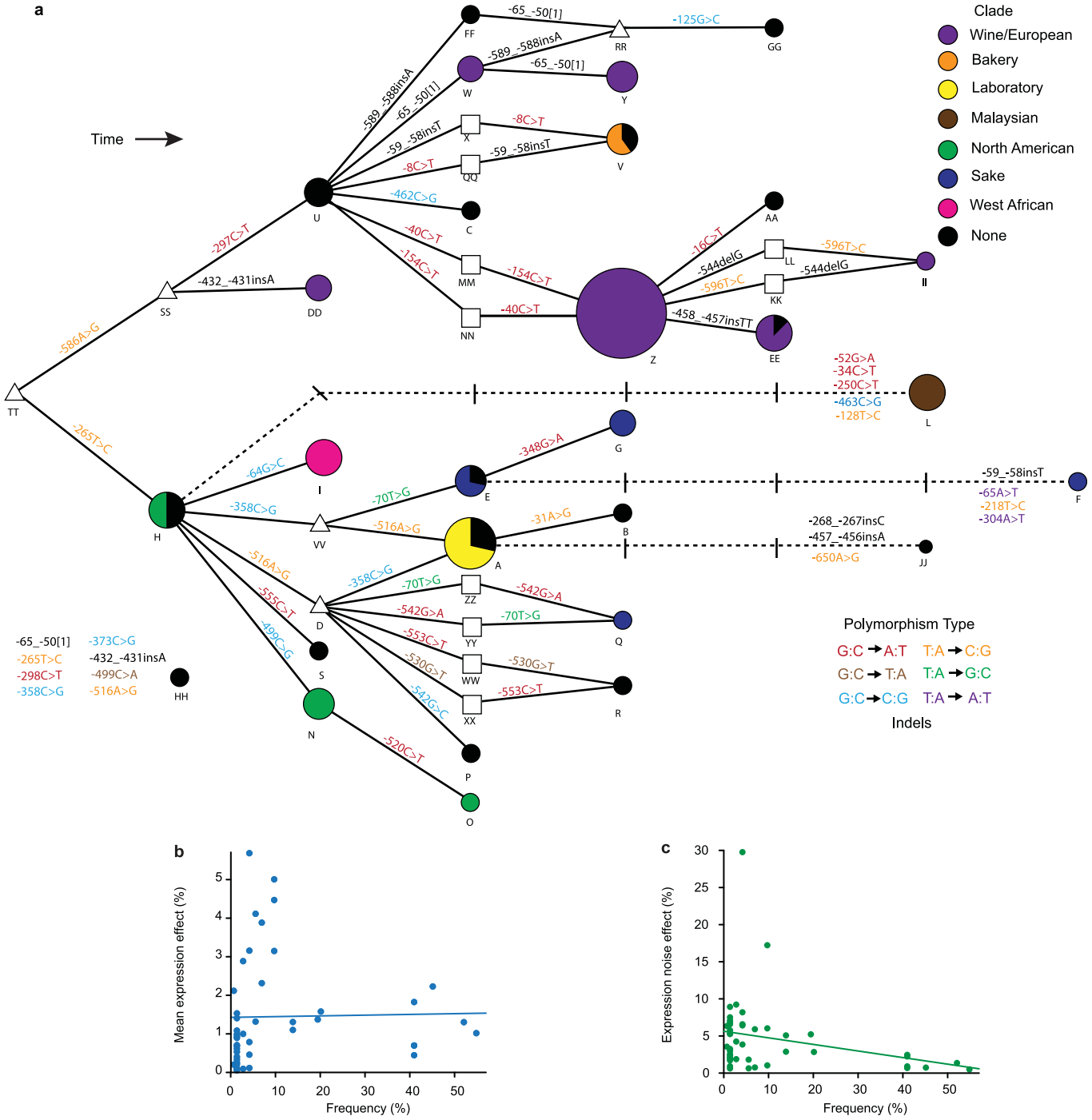
**Analysis of additional mutational data sets.** To test for differences in effects among different types of point mutation, we analysed data from previously published mutagenesis experiments in which the effects of individual mutations on *cis*-regulatory activity were determined[13–16]. Effects were split into each of the 12 mutation types and plotted on the same scale for all regulatory elements (Extended Data Fig. 3). For each *cis*-regulatory element, we used an ANOVA to test for a significant difference among mutation types. In all cases, no significant effect was observed ($P > 0.05$). We also used a linear model including the identity of the *cis*-regulatory element and mutation type as main effects to test for a significant difference among mutational classes for sets of *cis*-regulatory elements across studies. Again, we found no significant difference among different types of mutation ($P = 0.68$, ANOVA).


30. Taly, J.-F. *et al.* Using the T-Coffee package to build multiple sequence alignments of protein, RNA, DNA sequences and 3D structures. *Nature Protocols* **6,** 1669–1682 (2011).
31. Löytynoja, A. & Goldman, N. webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinform.* **11,** 579 (2010).
32. Libkind, D. *et al.* Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc. Natl Acad. Sci. USA* **108,** 14539–14544 (2011).
33. Scannell, D. R. *et al.* The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3* **1,** 11–25 (2011).
34. Liti, G. *et al.* High quality *de novo* sequencing and assembly of the *Saccharomyces arboricolus* genome. *BMC Genom.* **14,** 69 (2013).
35. Wang, Q.-M., Liu, W.-Q., Liti, G., Wang, S.-A. & Bai, F.-Y. Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. *Mol. Ecol.* **21,** 5404–5417 (2012).
36. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30,** 2725–2729 (2013).
37. Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9,** 1657–1659 (2000).
38. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38,** W529–W533 (2010).
39. Hittinger, C. T. *Saccharomyces* diversity and evolution: a budding model genus. *Trends Genet.* **29,** 309–317 (2013).
40. Wittkopp, P. J., Haerum, B. K. & Clark, A. G. Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430,** 85–88 (2004).
41. Wittkopp, P. J. in *Molecular Methods for Evolutionary Genetics* Vol. 772 (eds Orgogozo, V. & Rockman, M. V.) 297–317 (Humana, 2011).
42. Gietz, R. & Woods, R. in *Methods in Molecular Biology* 2nd edn, Vol. 313 (ed. Xiao, W.) 107–120 (Springer, 2006).

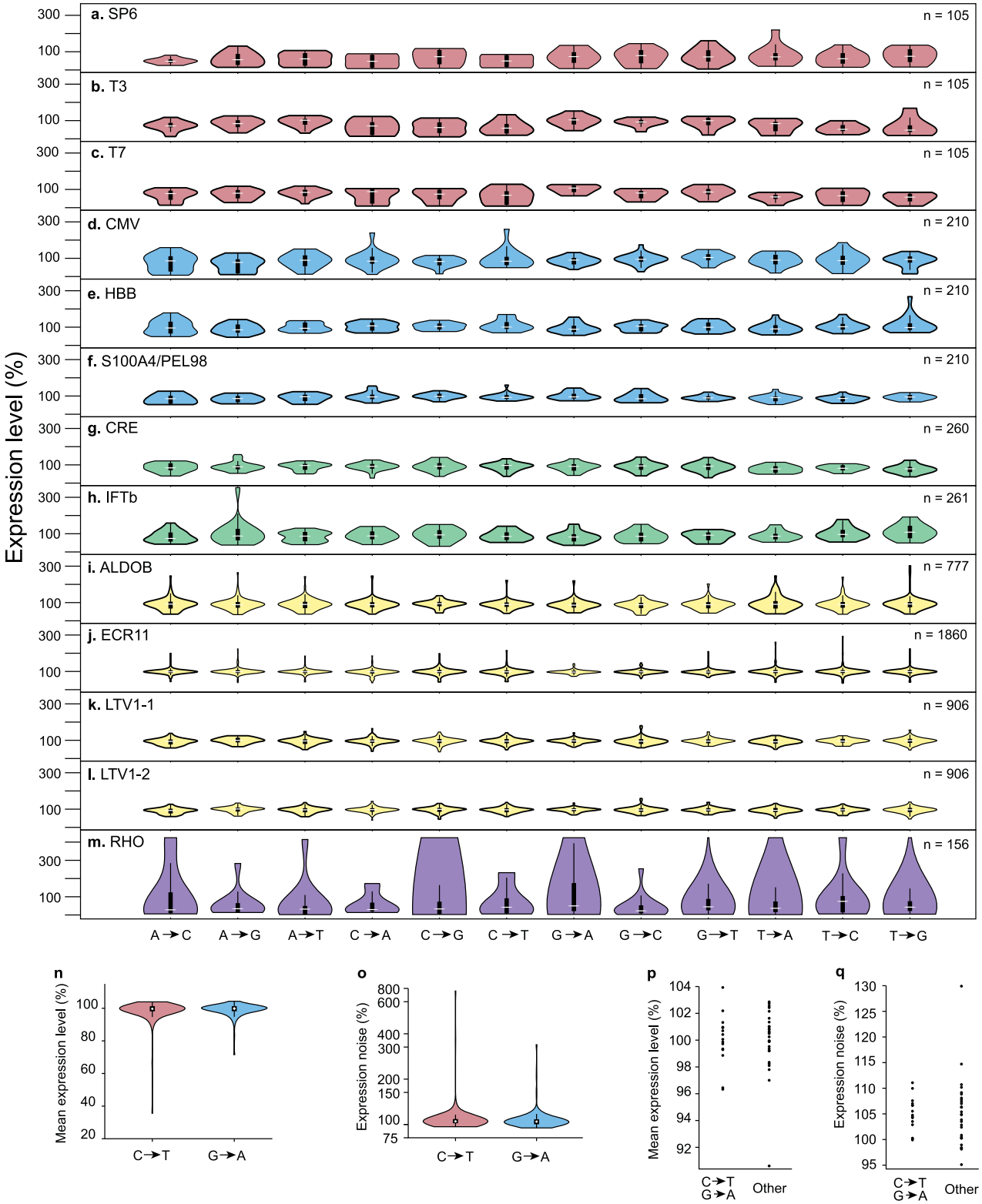
43. Kudla, G., Murray, A., Tollervey, D. & Plotkin, J. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324,** 255–258 (2009).
44. Lo, K., Hahne, F., Brinkman, R. R. & Gottardo, R. flowClust: a Bioconductor package for automated gating of flow cytometry data. *BMC Bioinform.* **10,** 145 (2009).
45. Hahne, F. *et al.* flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinform.* **10,** 106 (2009).
46. R Core Team. R: A language and environment for statistical computing (2013).



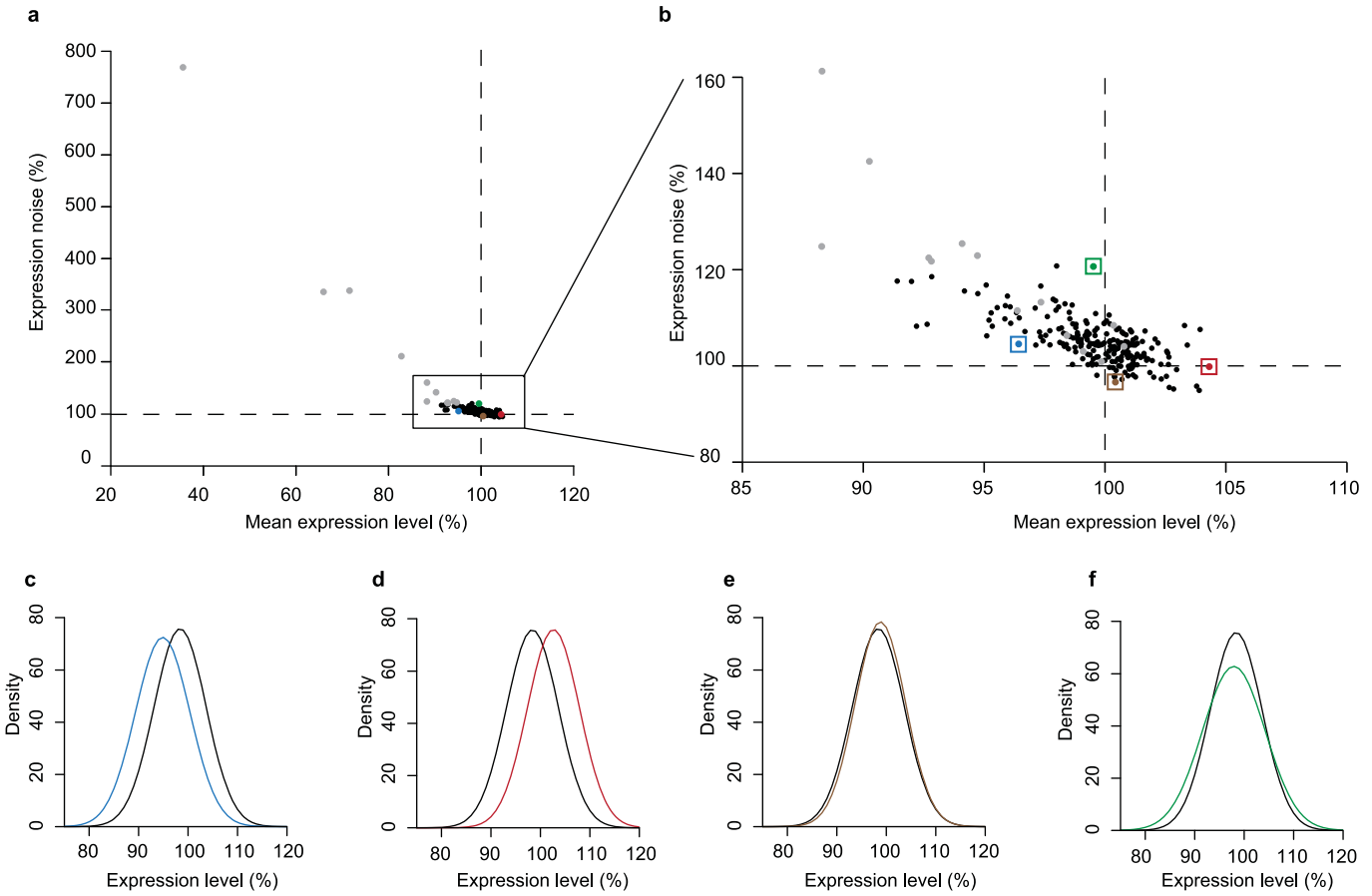

**Extended Data Figure 1 | *TDH3* promoter polymorphisms influence *TDH3* mRNA levels.** **a**, Locations of polymorphisms within the *TDH3* promoter relative to known functional elements, including RAP1 and GCR1 transcription factor binding sites, are shown. Squares, point mutations; circles, indels. Red, G:C→A:T; yellow, G:C→T:A; blue, G:C→C:G; orange T:A→C:G; green, T:A→G:C; purple, T:A→A:T. **b**, The $\log_2$ ratio of total expression divergence between natural isolates and a reference strain (*x* axis) versus the $\log_2$ ratio of total *cis*-regulatory expression divergence between natural isolates and the reference strain (*y* axis). Error bars, 95% confidence intervals. The 25 of 48 strains with significant *cis*-regulatory differences from the reference strain are shown in blue. Reference strain is shown in red. These data show differences in *cis*- and *trans*-regulation among strains, but do not reveal the evolutionary changes that give rise to these differences.



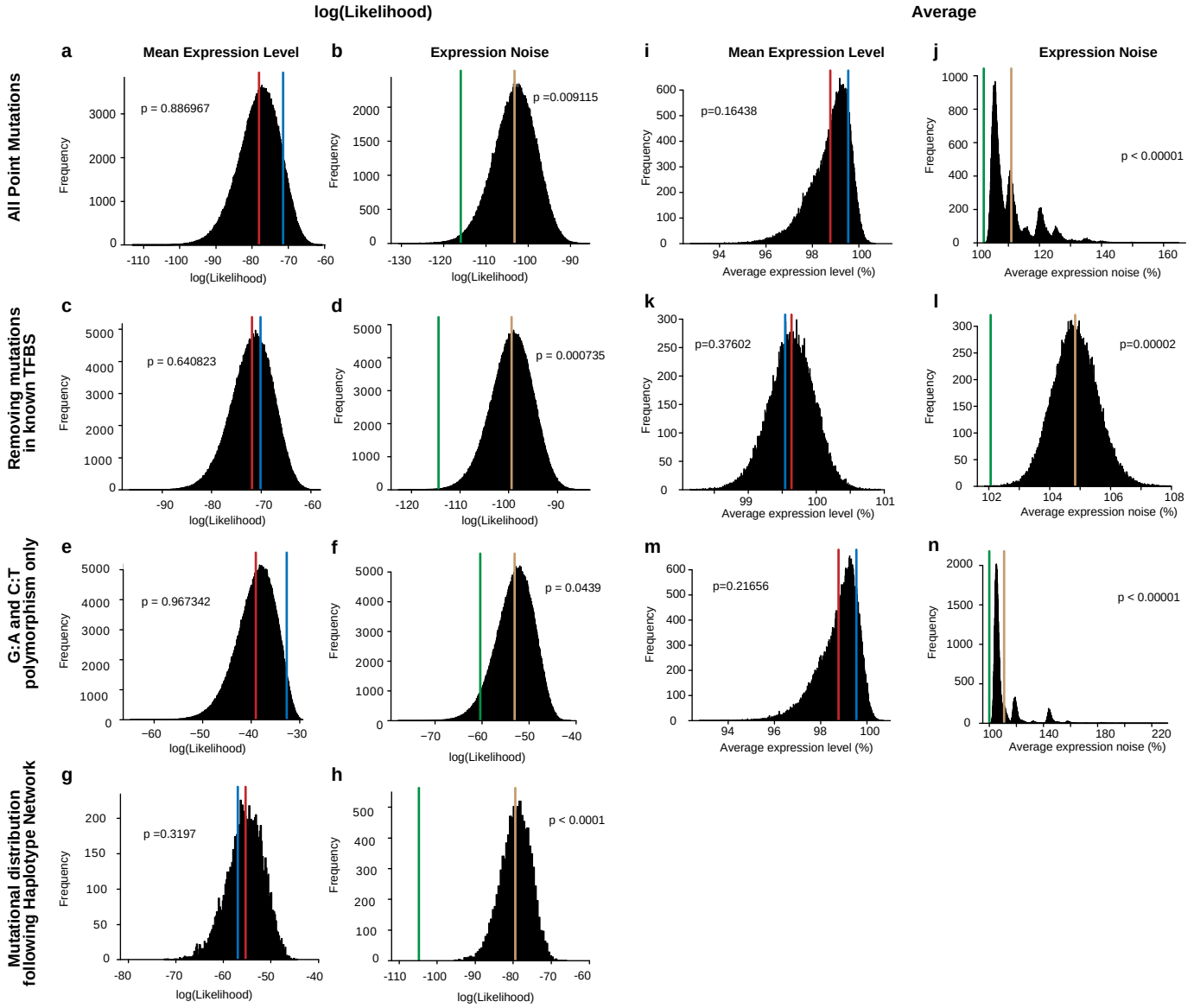

**Extended Data Figure 2 | Ancestral state reconstruction of the TDH3 promoter.** **a**, The *TDH3* promoter haplotype network is shown with the inferred ancestral strain at the left. Circles represent haplotypes observed among the 85 strains, with their diameters proportional to haplotype frequency. The haplotypes are coloured according to clade (Supplementary Table 1). Triangles are haplotypes that were not observed among the strains sampled, but must exist or have existed as intermediates between observed haplotypes. Squares are possible intermediates connecting two observed haplotypes, but it is unknown which of these actually exists or existed in *S. cerevisiae*. Solid lines connect haplotypes that differ by a single mutation; dashed lines connect haplotypes that differ by multiple mutations. Mutations on each branch are coloured by the mutation type as in Extended Data Fig. 1a. **b**, Relationship between the effect of a polymorphism on mean expression level and the frequency of that polymorphism among the strains sampled ($P = 0.43$). **c**, Relationship between the effect of a polymorphism on expression noise and the frequency of that polymorphism among the strains sampled ($P = 0.0028$).



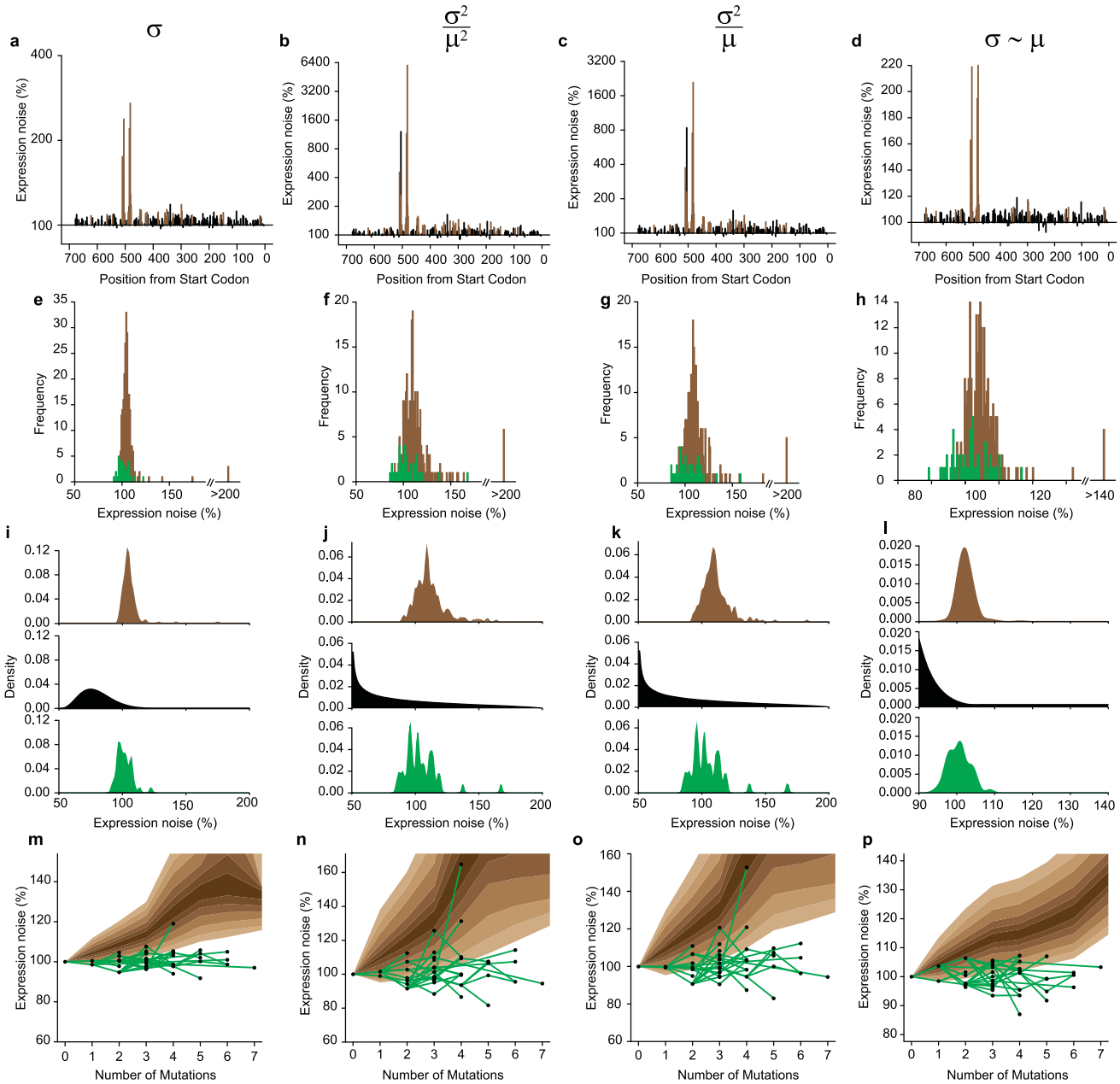

**Extended Data Figure 3 | No significant difference between mutation types.** Distributions of effects on mean expression level from previous random mutagenesis experiments are shown partitioned by mutation type. For each mutation type, the distribution (inside) and density (outside, coloured) of the effects on mean expression level are shown. The number of mutations tested for each promoter is shown in the upper right corner of each panel. **a**, Bacteriophage SP6 promoter. **b**, Bacteriophage T3 promoter. **c**, Bacteriophage T7 promoter. **d**, Human CMV promoter. **e**, Human HBB promoter. **f**, Human S100A4/PEL98 promoter. **g**, Synthetic cAMP-regulated enhancer. **h**, Interferon-β enhancer. **i**, ALDOB enhancer. **j**, ECR11 enhancer. **k**, LTV1 enhancer replicate 1. **l**, LTV1 enhancer replicate 2. **m**, Rhodopsin promoter. Red: bacteriophage promoters from ref. 13. Blue: mammalian promoters from ref. 13. Green: mammalian enhancers from ref. 14. Yellow: mammalian promoters from ref. 15. Purple: promoter from ref. 16. **n**, Distribution of effects for C→T (red) and G→A (blue) mutations for mean expression level in this study. **o**, Same as **n**, but for expression noise. **p**, Distribution of effects for C→T/G→A polymorphisms compared with other polymorphism types for mean expression level in this study. **q**, Same as **p**, but for gene expression noise.




**Extended Data Figure 4 | Correlation between mean expression level and expression noise.** **a**, Correlation between mean expression level (*x* axis) and expression noise (*y* axis) for the 236 point mutations in the *TDH3* promoter ($R^2 = 0.85$). Grey points correspond to mutations in known transcription factor binding sites. Coloured points correspond to individual mutations highlighted in **c**–**f**. **b**, Alternative plot showing the majority of data from **a** more clearly; grey and coloured points are the same as in **a**. **c**, Distribution of gene expression phenotypes from a mutant (blue) with decreased mean expression level but similar expression noise as the reference strain (black). Outside the known TFBS, 50% of mutations decreased mean expression. **d**, Distribution of gene expression phenotypes from a mutant (red) with increased mean expression level but similar gene expression noise as the reference strain (black). Outside the known TFBS, 50% of mutations increased mean expression. **e**, Distribution of gene expression phenotypes from a mutant (brown) with decreased gene expression noise but similar mean expression level as the reference strain (black). Outside the known TFBS, 13% of mutations decreased expression noise. **f**, Distribution of gene expression phenotypes from a mutant (green) with increased gene expression noise but similar mean expression level as the reference strain (black). Outside the known TFBS, 87% of mutations increased expression noise.




**Extended Data Figure 5 | Tests for selection.** **a–h**, Tests for selection using likelihood. **a**, The distribution of likelihood values for 100,000 randomly sampled sets of 45 mutations drawn from the mutational effect distribution is shown for mean expression level. The average likelihood for all samples of mutations tested (red) as well as the likelihood of the observed polymorphisms (blue) are also shown. **b**, Same as **a**, but for expression noise. The average likelihood for all mutation samples tested is shown in brown and the likelihood of the observed polymorphisms is shown in green. **c**, Same as **a**, but with the large effect mutations in the TFBS removed from the mutational effect distribution used for sampling. **d**, Same as **b**, but after removing the mutations in the TFBS from the mutational effect distribution. **e**, Same as **a**, but using only G→A and C→T polymorphisms. **f**, same as **b**, but using only G→A and C→T polymorphisms. **g**, Distribution of likelihoods for 10,000 random walks along the *TDH3* promoter haplotype network using the effects from the mutational distribution. **h**, Same as **e**, but for expression noise. **i–n**, Tests for selection using average effects. **i**, The distribution of average effects for 100,000 randomly sampled sets of 45 mutations drawn from the mutational effect distribution is shown for mean expression level (black). Polymorphisms do not have a significantly different average mean expression (blue, 99.5%) than sets of mutations (red, 98.8%; $P = 0.16438$). This figure is comparable to Extended Data Fig. 5a, but uses average effects instead of the likelihoods to test for differences in distribution between random mutations and polymorphisms. **j**, Same as **i**, but for expression noise. Polymorphisms have significantly lower average expression noise (green, 102.1%) than sets of random mutations (brown, 110.9%; $P < 0.00001$). **k**, Same as **i**, but with the large effect mutations in the TFBS removed from the mutational effect distribution used for sampling (polymorphisms, 99.5%; mutations, 99.6%; $P = 0.37602$). **l**, Same as **j**, but after removing the mutations in the TFBS from the mutational effect distribution (polymorphisms, 102.1%; mutations, 104.8%; $P = 0.00002$). **m**, Same as **i**, but using only G→A and C→T polymorphisms (polymorphisms, 99.7%; mutations, 98.8%; $P = 0.21656$). **n**, Same as **j**, but using only G→A and C→T polymorphisms (polymorphisms, 100.0%; mutations, 110.9%; $P < 0.00001$).

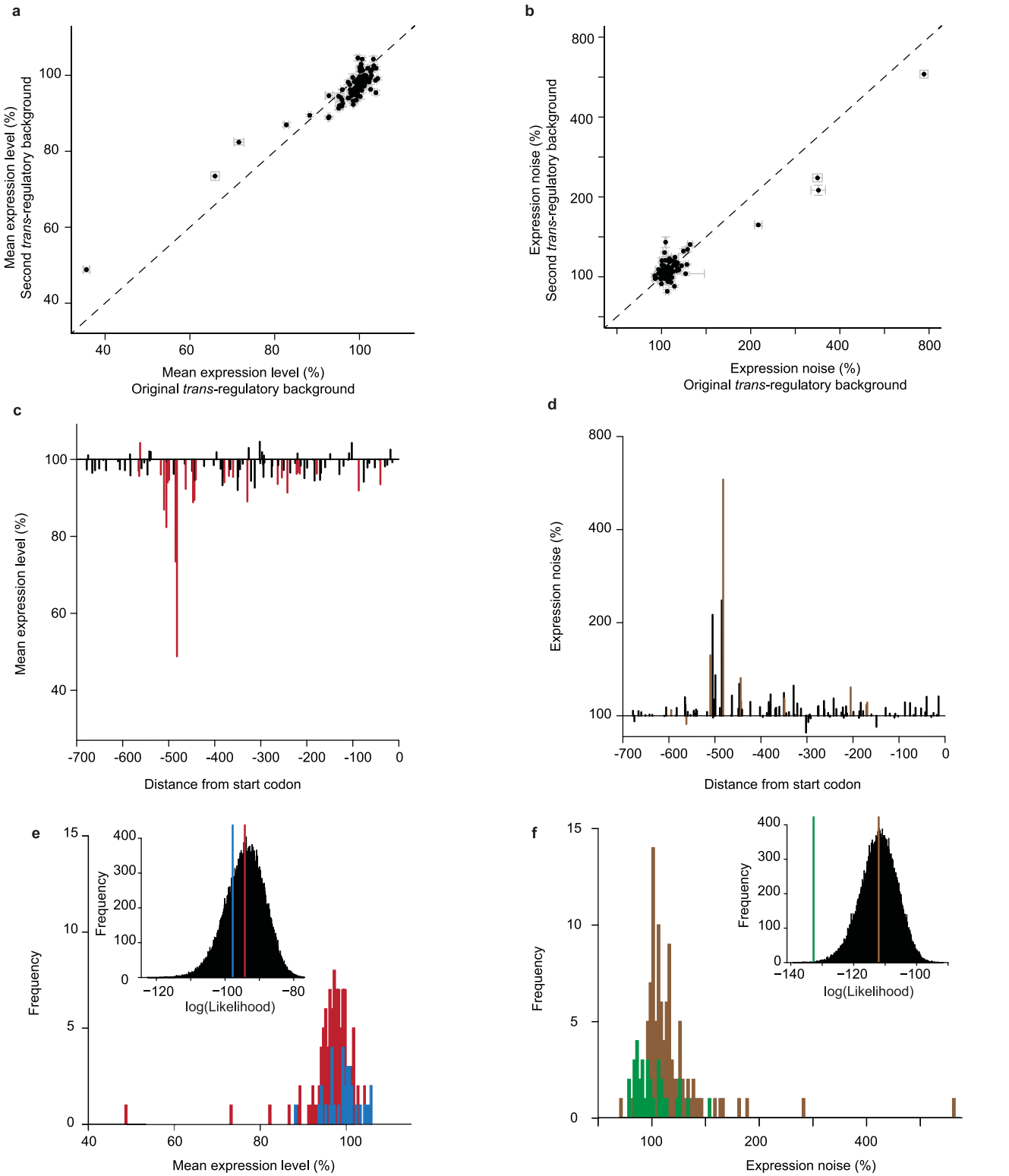



**q**

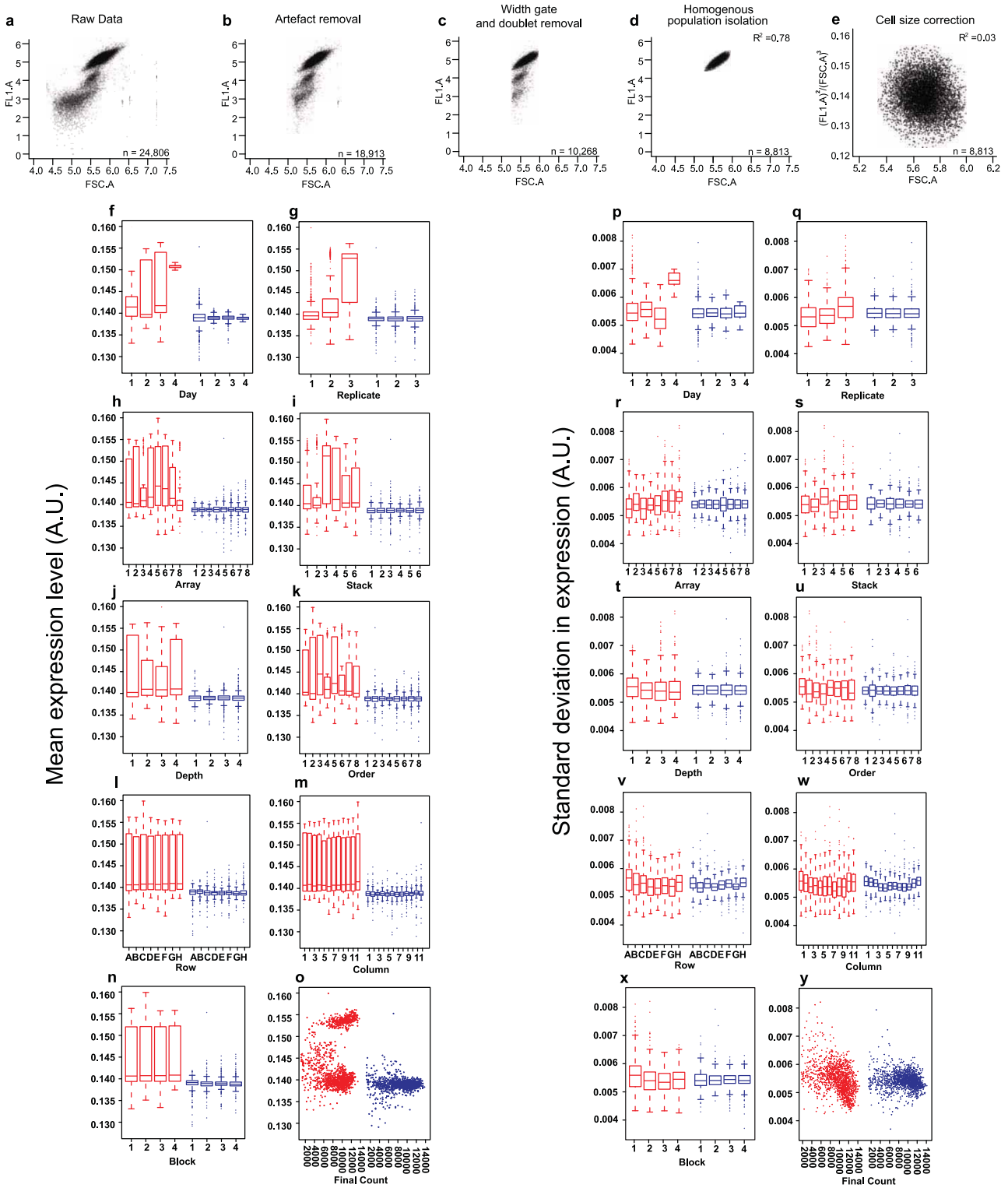
| Metric*† | Distribution§ | No TFBS Mutations§ | G:A/C:T Polymorphism§ | Fitness Function¶ | Network‖ |
|---|---|---|---|---|---|
| μ | 0.88697 | 0.64082 | 0.96734 | 0.87 | 0.3197 |
| σ | <0.00001 | <0.00001 | <0.00001 | $9x10^{-6}$ | <0.0001 |
| σ/μ | 0.00912 | 0.00076 | 0.04390 | 0.00019 | <0.0001 |
| $σ^2/μ^2$ | <0.00001 | <0.00001 | 0.00760 | 0.00002 | <0.0001 |
| $σ^2/μ$ | <0.00001 | <0.00001 | <0.00001 | 0.00015 | <0.0001 |
| Regression‡ | <0.00001 | <0.00001 | <0.00001 | $4x10^{-7}$ | <0.0001 |

*mean expression (μ) †standard deviation in expression (σ) ‡Residuals from linear model: σ ~ μ
§100,000 permutations ‖10,000 permutations ¶Likelihood Ratio

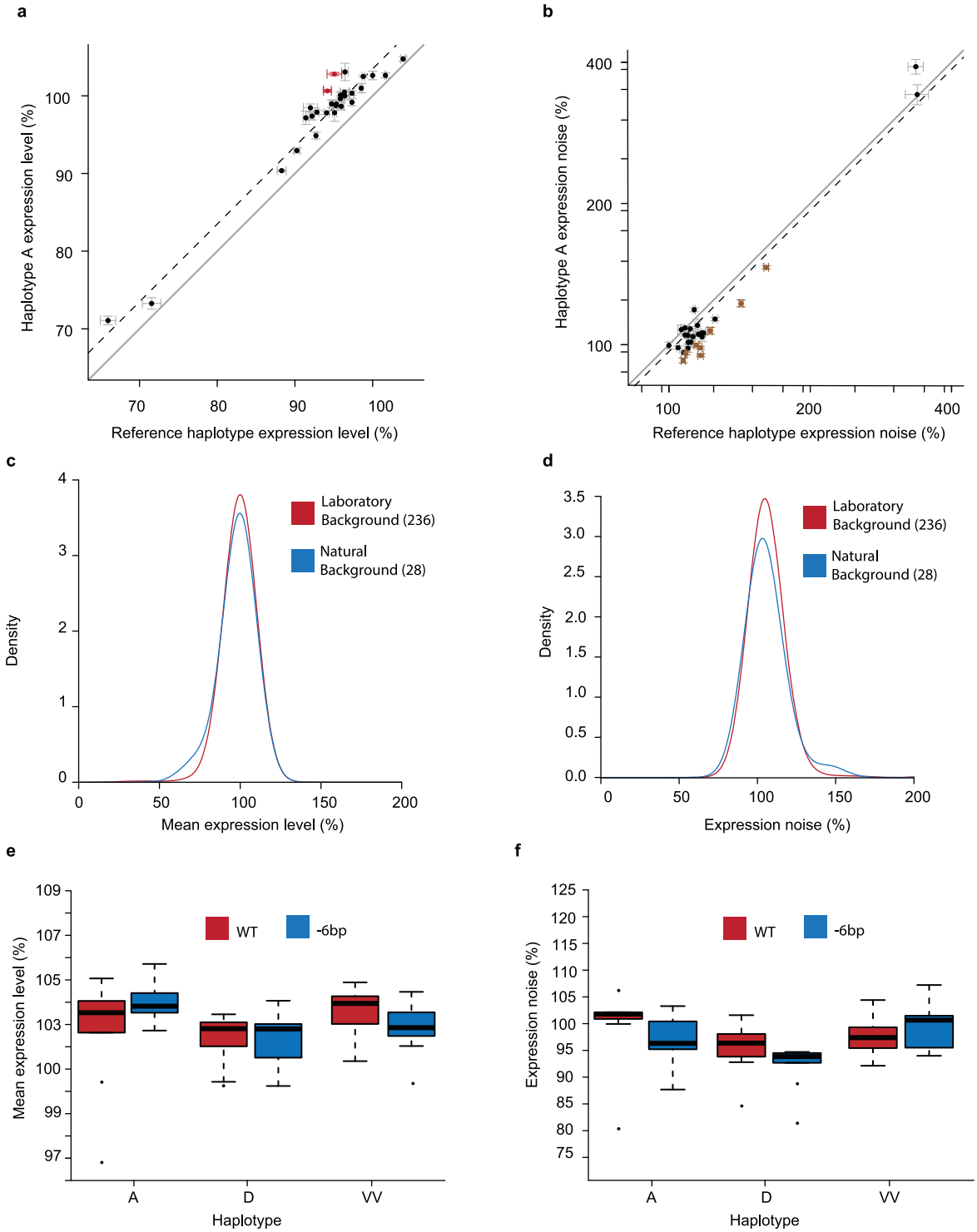
**Extended Data Figure 6 | Test for selection using alternative metrics for quantifying gene expression noise.** **a–d**, Distributions of effects for mutations on gene expression noise across the *TDH3* promoter with expression noise quantified as $\sigma$ (**a**), $\sigma^2/\mu^2$ (**b**), $\sigma^2/\mu$ (**c**), and residuals from the regression of $\sigma$ on $\mu$ (**d**). **e–h**, Distributions of effects for mutations on gene expression noise (brown) compared with polymorphisms (green) with noise quantified as $\sigma$ (**e**), $\sigma^2/\mu^2$ (**f**), $\sigma^2/\mu$ (**g**), and residuals from the regression of $\sigma$ on $\mu$ (**h**). **i–l**, The maximum likelihood fitness function (middle, black) relating the distribution of mutational effects (top, brown) to the distribution of observed polymorphisms (bottom, green) for expression noise quantified as $\sigma$ (**i**), $\sigma^2/\mu^2$ (**j**), $\sigma^2/\mu$ (**k**), and residuals from the regression of $\sigma$ on $\mu$ (**l**). **m–p**, Changes in expression noise observed among haplotypes over time in the inferred haplotype network (Extended Data Fig. 2a) are shown in green. The brown background represents the 95th, 90th, 80th, 70th, 60th, and 50th percentiles, from light to dark, for expression noise resulting from 10,000 independent simulations of phenotypic trajectories in the absence of selection where noise is quantified as $\sigma$ (**m**), $\sigma^2/\mu^2$ (**n**), $\sigma^2/\mu$ (**o**), and residuals from the regression of $\sigma$ on $\mu$ (**p**). **q**, *P* values for tests of selection using mean expression ($\mu$) and five metrics of expression noise, including $\sigma/\mu$ which is used throughout the main text.



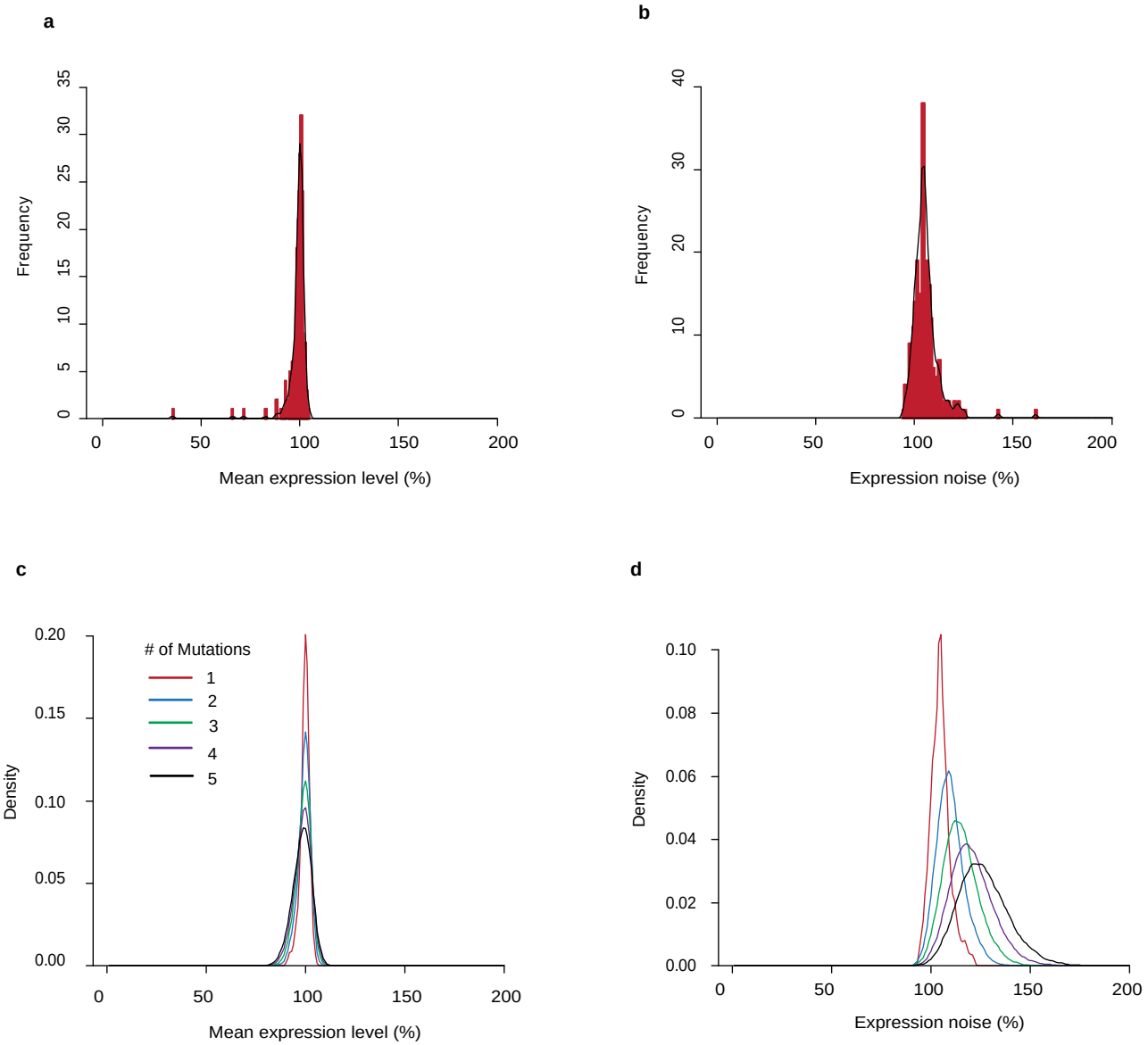

**Extended Data Figure 7 | Effects of mutations and polymorphisms on a second *trans*-regulatory background.** **a**, A comparison between effects of mutations on mean expression in the original *trans*-regulatory background (*x* axis) and a hybrid *trans*-regulatory background between BY4741 and YPS1000 (*y* axis). Error bars, 95% confidence intervals. **b**, Same as **a**, but for gene expression noise. **c**, Effects of individual mutations on mean expression level in the hybrid *trans*-regulatory background are shown in terms of the percentage change relative to the un-mutagenized reference allele, and are plotted according to the site mutated in the 678 bp region (significant mutations: red lines, *t*-test, Bonferroni corrected). Note that most mutations decrease expression, unlike in the original genetic background. **d**, Same as **c**, but for gene expression noise (significant mutations: brown lines, *t*-test, Bonferroni corrected). **e**, Distribution of *de novo* mutation effects in the second *trans*-regulatory background (red) compared with the effects of naturally occurring haplotypes in this *trans*-regulatory background (blue). Inset: the distribution of likelihood values for 100,000 randomly sampled sets of 27 mutations drawn from the mutational effect distribution is shown for mean expression level. The average likelihood for all samples of mutations tested (red) as well as the likelihood of the observed polymorphisms (blue) are also shown ($P = 0.2584$). Removing mutations in the known TFBS resulted in a significant difference between mutations and polymorphisms ($P = 0.00781$). **f**, Same as **e**, but for gene expression noise. Mutations, brown. Polymorphisms, green ($P = 0.00037$). Removing mutations in the known TFBS did not change this result ($P < 0.00001$).




**Extended Data Figure 8 | Methodology for the analysis of flow cytometry data.** **a**, Raw data from the flow cytometer are shown for the first control sample collected. Each point is an individual event scored by the flow cytometer, the vast majority of which are expected to be cells. FSC.A is a proxy for cell size, and FL1.A is a measure of YFP fluorescence. $Log_{10}$ values are plotted both for FSC.A and for FL1.A. **b**, The same sample is shown after events found in the negative control sample (using hard gates on FSC.A and FL1.A) were excluded. **c**, The same sample is shown after flowClust was used to remove events likely to be from multiple cells entering the detector simultaneously. **d**, The same sample is shown after flowClust was used to isolate the densest homogenous population within the sample. The $R^2$ value shown is the correlation between YFP fluorescence and cell size. **e**, After correcting for differences in cell size, the correlation between YFP fluorescence and cell size was nearly 0 and not significant. In all panels, the number of events analysed (that is, sample size) is shown in the bottom right corner. Box plots of mean expression of control samples before (red) and after (blue) correcting for the effects of individual plates for each day on which samples were run (**f**), for replicates nested within day (**g**), for array nested within day and replicate (**h**), for stack nested within day (**i**), for depth nested within day (**j**), for order nested within day and replicate (**k**), for row nested within array (**l**), for column nested within array (**m**), for block nested within array (**n**), and for the final cell count (**o**). The *y* axis is in arbitrary units. **p**–**x**, Same as **f**–**o**, but for gene expression noise.




**Extended Data Figure 9 | Consistency of mutational effects on different genetic backgrounds.** **a**, The effects on mean expression level for each of the 28 mutations tested on both the reference haplotype (*x* axis) and natural haplotype A observed in wild strains (*y* axis) are shown. These two haplotypes differ by a single point mutation. Solid lines show expression from the $P_{TDH3}$ haplotypes on which the two sets of mutations were created, both of which were defined as 100% activity. The grey line shows $y = x$. The dashed line shows the consistent increase in mean expression level when these mutations were tested on haplotype A. Error bars, 95% confidence intervals. Coloured points have significantly different effects on the two backgrounds ($P < 0.05$, ANOVA, Bonferroni corrected), indicating weak epistasis. **b**, Same as **a**, but for gene expression noise. **c**, Distributions of mutational effects for mean expression levels based on the 236 point mutations tested on the reference haplotype (red) as well as for the 28 mutations tested on haplotype A (blue). **d**, Same as **c**, but for gene expression noise. **e**, The effect on mean expression of the full *TDH3* promoter (red) compared with promoters containing six fewer base pairs at the 5′ end (blue). Each box plot summarizes data from nine replicates. **f**, Same as **e**, but for expression noise.




**Extended Data Figure 10 | Probability distributions for mutational effects.** **a**, A histogram summarizing the mutational effects on mean expression level is shown (red), overlaid with the density curve (black line) used to calculate the likelihood of an effect on mean expression level. **b**, Same as **a**, but for expression noise. **c**. Density curves for the effects of one (red), two (blue), three (green), four (purple), or five (black) mutations randomly drawn from the distribution of mutational effects observed for mean expression level. **d**, Same as **c**, but for expression noise.