nature genetics

Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks

Xionglei He[1,3,4], Wenfeng Qian[1,4], Zhi Wang[1,4], Ying Li[1,2] & Jianzhi Zhang[1]

Epistasis refers to the interaction between genes. Although high-throughput epistasis data from model organisms are being generated and used to construct genetic networks[1–3], the extent to which genetic epistasis reflects biologically meaningful interactions remains unclear[4–6]. We have addressed this question through *in silico* mapping of positive and negative epistatic interactions amongst biochemical reactions within the metabolic networks of *Escherichia coli* and *Saccharomyces cerevisiae* using flux balance analysis. We found that negative epistasis occurs mainly between nonessential reactions with overlapping functions, whereas positive epistasis usually involves essential reactions, is highly abundant and, unexpectedly, often occurs between reactions without overlapping functions. We offer mechanistic explanations of these findings and experimentally validate them for 61 *S. cerevisiae* gene pairs.

Epistasis refers to the phenomenon in which the effect of a gene on a trait is masked or enhanced by one or more other genes[6,7]. Fisher and other population and quantitative geneticists extended the concept to mean nonindependent or nonmultiplicative effects of genes[6,8]. The direction, magnitude and prevalence of epistasis is important for understanding gene function and interaction[2,6,9], speciation[10], evolution of sex and recombination[11,12], evolution of ploidy[13], mutation load[14], genetic buffering[15], human disease[4,5] and drug-drug interaction[16]. Epistasis in fitness between two mutations is commonly defined by $\varepsilon = W_{XY} - W_X W_Y$, where W_X and W_Y represent the fitness values of two single mutants relative to the wild-type, and W_{XY} represents the fitness of the corresponding double mutant. Epistasis is said to be positive when $\varepsilon > 0$ and negative when $\varepsilon < 0$. Where deleterious mutations are concerned, positive epistasis lessens the fitness reduction predicted from individual mutational effects, whereas negative epistasis enhances it. The magnitude of epistasis between different pairs of mutations may be compared using scaled epistasis $\tilde{\varepsilon}$ (ref. 17), which is transformed from and has the same sign as ε but is normally bounded by the values −1 and 1. We apply flux balance analysis (FBA) of metabolic networks[18] to explore the functional association between biochemical reactions that are epistatic to each other. Assuming a steady state in metabolism, FBA maximizes the rate of biomass production under the stoichiometric matrix of all reactions and a set of flux constraints. The maximized rate in a mutant strain relative to that in the wild-type strain can be regarded as the fitness of the mutant relative to the wild-type[17]. FBA can be used to investigate the fitness of the cell under various environmental and genetic perturbations[19,20] and has been used to generate the epistasis map of yeast metabolic genes[17,21,22]. We first study the bacterium *Escherichia coli*, because its reconstructed metabolic network is of high quality and its FBA predictions have been empirically verified[20,23].

Using FBA, we identified from the *E. coli* metabolic network 270 reactions whose removal reduces the organism's fitness in glucose

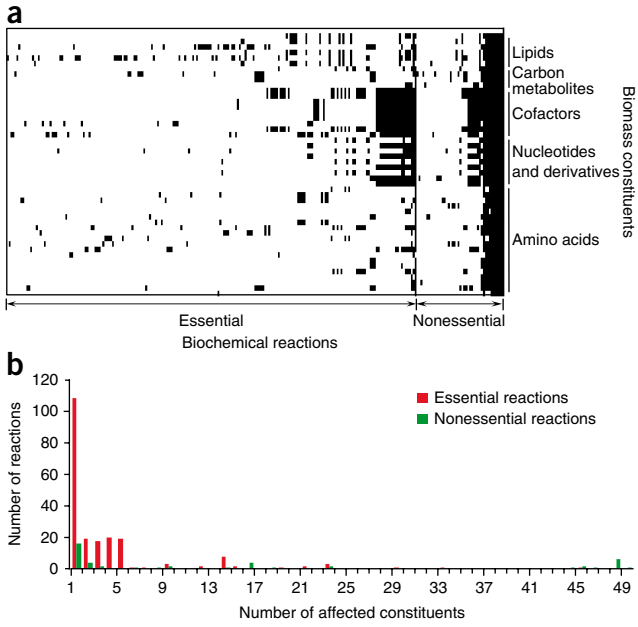

Figure 1 Functions of *E. coli* metabolic reactions in glucose minimal medium. (**a**) Functions of 255 important reactions in producing 49 biomass constituents. Columns represent reactions and rows represent biomass constituents. (**b**) Distribution of the number of biomass constituents affected by a reaction.

[1]Department of Ecology and Evolutionary Biology, Ann Arbor, Michigan, USA. [2]Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China. [3]Present address: State Key Laboratory of Bio-control, College of Life Sciences, Sun Yat-Sen University, Guangzhou, China. [4]These authors contributed equally to this work. Correspondence should be addressed to J.Z. (jianzhi@umich.edu).

Received 21 July 2009; accepted 18 December 2009; published online 24 January 2010; doi:10.1038/ng.524

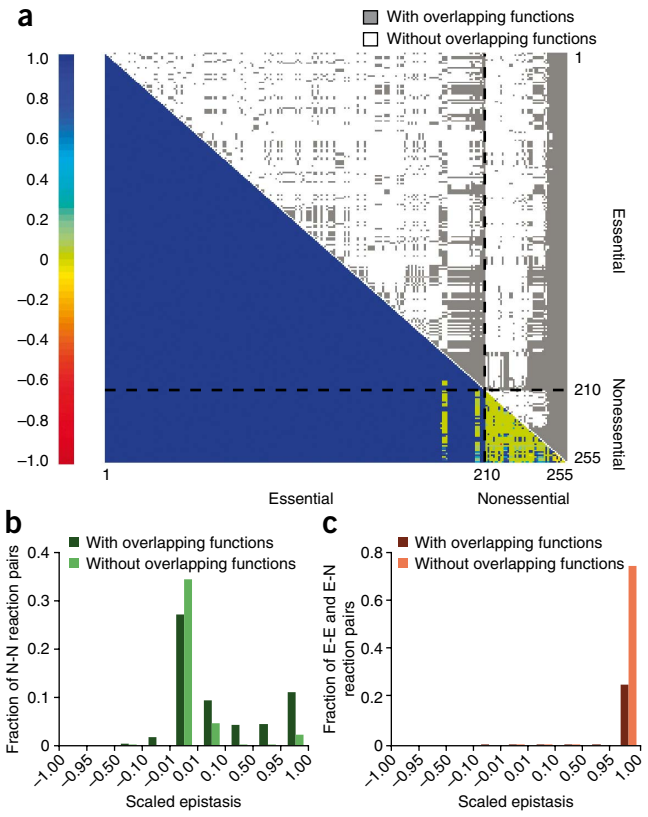

Figure 2 Pairwise epistasis and functional association among 255 important reactions in *E. coli*. (**a**) An overview of epistasis and functional association among reactions. Both rows and columns represent reactions. Scaled epistasis between reactions is shown in the lower left triangle by the heat map. Functional association between reactions is presented in the upper right triangle, where a gray dot is shown when two reactions have overlapping functions. Epistasis and reaction functions are both determined in the glucose minimal medium. (**b**) Frequency distribution of scaled epistasis between nonessential reactions. (**c**) Frequency distribution of scaled epistasis between two reactions that include at least one essential reaction. E, essential; N, nonessential. Note the difference in *y* scale between **b** and **c**.

minimal medium. Removing any of the remaining 661 reactions has no such effect, usually because the reaction has zero flux in this medium, and occasionally because the network includes another reaction that can fully compensate its loss. Among the 270 reactions, 212 are essential, meaning that deleting any one of them results in zero fitness. We considered a genetic perturbation in each reaction that constrains its flux to ≤50% of its wild-type optimal value and then computed the fitness of the mutant by FBA. We similarly computed the fitness values of all possible double mutants and obtained ε and $\tilde{\varepsilon}$ for all pairs of the 270 reactions, which revealed the global epistasis pattern in the metabolic network (**Supplementary Table 1**). Constraining the flux to ≤50% instead of zero[17,21,22] allowed us to investigate essential reactions. Consequently, the number of pairwise epistasis values obtained here exceeds 25 times that previously obtained[17]. Constraining the flux to other nonzero levels did not alter our results qualitatively (**Supplementary Table 1**).

To examine whether metabolic reactions with epistatic relationships are functionally associated, we needed to identify the function of each reaction in generating the *E. coli* biomass, which is composed of 49 constituents. If a reaction is important for producing a set of biomass constituents, the removal of these constituents from the biomass function will restore the biomass that was reduced as a result of the deletion of the reaction. On the basis of this idea, we designed a removal-recovery method to determine the functions of 255 of the 270 important reactions in generating biomass constituents (**Fig. 1a**). For the remaining 15 reactions, the functions cannot be unambiguously determined, and thus they were excluded from our analysis. Most of the 255 reactions each contribute to only one biomass constituent, whereas a few reactions affect many or even all 49 constituents (**Fig. 1b**). Note that we again used glucose minimal medium in determining the function of each reaction, because some reactions have variable functions in different media. Functional assignment by our method was generally consistent with the conventional functional annotation of *E. coli* reactions[24], but our assignment is expected to be more precise in identifying the biomass constituents contributed by each reaction.

We found 26 (0.08%) reaction pairs that show apparent negative epistasis ($\tilde{\varepsilon} \leq -0.01$). Among them, 25 pairs each share functions in producing at least one biomass constituent (**Fig. 2a**,**b** and **Table 1**). The remaining pair comprises the reactions catalyzed by malate synthase (MALS) and phosphoenolpyruvate carboxylase (PPC), anaplerotic reactions that feed the Krebs cycle. The lack of shared biomass constituents between them is due to the incomplete identification of MALS and PPC functions caused by their mutual functional compensation (**Supplementary Fig. 1**). A common interpretation of negative epistasis between two genes is that the two genes can individually perform a common function, and therefore each can compensate for the loss of the other. Our observation that virtually every pair of reactions with negative epistasis shares at least one function strongly support this interpretation (**Fig. 2b**). Although negative epistasis might be expected to occur between two nonessential reactions, this is not absolute. For example, two essential reactions (or one essential reaction and one nonessential reaction) may share a

Table 1 Numbers of reaction pairs that show epistatic relationships in glucose minimal medium

		Epistasis in *E. coli*[2]			Epistasis in yeast[2]		
Reaction pairs[1]	Functions	Negative	Zero	Positive	Negative	Zero	Positive
E-E	With overlap	0	9	4,269	1	2	1,780
	Without overlap	0	0	17,667	0	3	10,617
	Sum	0	9	21,936	1	5	12,397
E-N	With overlap	3	83	3,704	10	67	2,153
	Without overlap	0	34	5,626	0	99	6,203
	Sum	3	117	9,330	10	166	8,356
N-N	With overlap	22	267	288	24	137	402
	Without overlap	1	339	73	7	661	200
	Sum	23	606	361	31	798	602
All	With overlap	25	359	8,261	35	206	4,335
	Without overlap	1	373	23,366	7	763	17,020
	Sum	26	732	31,627	42	969	21,355

[1]Pairwise relationships among 255 important *E. coli* reactions and among 212 important yeast reactions. E, essential reaction; N, nonessential reaction. [2]Scaled epistasis of ≥0.01 is considered positive, ≤−0.01 is considered negative and between −0.01 and 0.01 is considered zero.

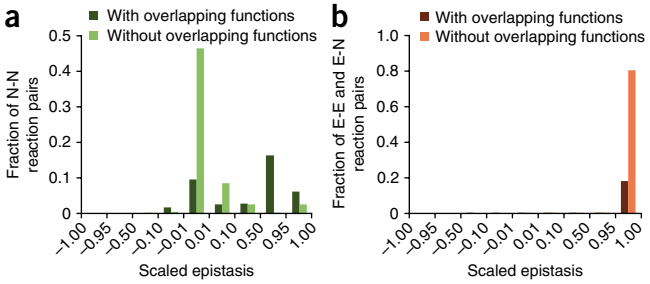

Figure 3 Pairwise epistasis and functional association among 212 important reactions in yeast. (**a**) Frequency distribution of scaled epistasis between nonessential reactions. (**b**) Frequency distribution of scaled epistasis between two reactions that include at least one essential reaction. E, essential; N, nonessential. Note the difference in *y* scale between **a** and **b**.

nonessential function in producing a biomass constituent and show negative epistasis by this common function (**Table 1**).

In contrast to the rarity of negative epistasis, >97% of reaction pairs show apparent positive epistasis ($\tilde{\varepsilon} \geq 0.01$) (**Fig. 2a**). However, only ~26% of these occur between reactions that share at least one biomass constituent (**Fig. 2c** and **Table 1**). There is also no significant difference in ε or $\tilde{\varepsilon}$ between functionally overlapping and nonoverlapping reaction pairs with positive epistasis. It is often observed that a reaction is positively epistatic with many apparently unrelated reactions. Use of ε instead of $\tilde{\varepsilon}$ in measuring epistasis does not change this pattern. The lack of functional overlap between most positively epistatic reaction pairs challenges the general interpretation of epistasis as reflecting functional association[2,9,25].

Why does positive epistasis occur so frequently between functionally unrelated reactions? **Figure 2a** shows that virtually every essential reaction shows strong positive epistasis ($\tilde{\varepsilon} \sim 1$) with any other reaction, regardless of its function and essentiality. This can be explained by considering that, when an essential reaction is constrained, almost all other reactions in the network do not work in their full capacity, such that the composition stoichiometry of the biomass is still maintained (**Supplementary Fig. 2a**,**b**). Consequently, a genetic perturbation in a second reaction that reduces its capacity will have a negligible additional effect, resulting in positive epistasis. Note that positive epistasis sometimes occurs between nonessential genes, and in these cases ~80% (288 of 361) show functional overlaps (**Fig. 2b**).

Why is there no such effect between nonessential reactions? There are three requirements for a metabolic reaction to be considered here as important but nonessential. First, it must function in producing one or more biomass constituents. Second, there must be alternative reactions that can also make its product. Third, compared with the alternative reactions, it must be more efficient in producing at least one constituent. When the flux of a nonessential reaction is constrained, its less efficient alternative reaction will be turned on (**Supplementary Fig. 2c**). Owing to the lower efficiency of the alternative reaction, nutrients that previously went through other reactions for making other biomass constituents can be redistributed in such a way that the biomass reduction by the flux constraint is minimized (**Supplementary Fig. 2c**). It can be shown mathematically that when the number of reactions in the network is large, perturbations of two functionally unrelated nonessential reactions will have a nearly multiplicative effect on biomass production and cause negligibly weak positive epistasis[15,17] (**Supplementary Note**).

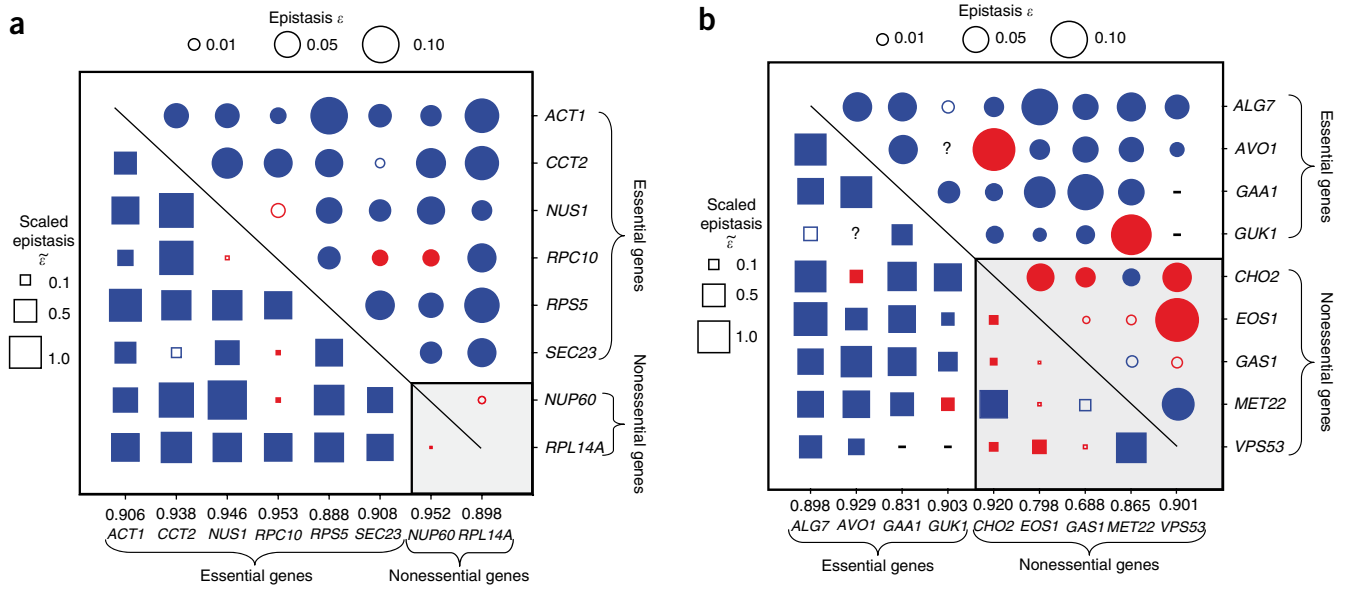

Figure 4 Epistasis (ε) and scaled epistasis ($\tilde{\varepsilon}$) among 17 yeast genes. Circles show ε and squares $\tilde{\varepsilon}$. Blue and red indicate positive and negative epistasis, respectively, whereas the areas of the circles and squares are proportional to the absolute values of ε and $\tilde{\varepsilon}$, respectively, with the scales given on the top and left sides of each panel. Solid symbols indicate statistically significant epistasis ($P < 0.05$), whereas open symbols indicate insignificant epistasis. The shaded area in the lower right corner shows relationships between nonessential genes. Fitness values of strains with genes replaced or inserted by *LEU2*, relative to the wild types, are presented on the *x* axis. (**a**) Epistasis among eight haploinsufficient genes, measured in diploid cells after deletion of one allele per gene. All genes belong to different functional categories with the exception of *RPS5* and *RPL14A*, both of which encode ribosomal proteins. (**b**) Epistasis among nine haplosufficient genes, measured in haploid cells after reduction of protein expression of essential genes and deletion of nonessential genes. All genes belong to different functional categories, with the exception of *GAA1* and *GAS1*. *MET22* and *CHO2* are metabolic genes, with FBA-predicted scaled epistasis equal to 1. '–' indicates that double-perturbation cells could not be obtained, probably because of unsuccessful experiments or synthetic lethality. '?' indicates that epistasis could not be measured owing to the lack of fitness effect of single perturbations with the *URA3* marker. In **Supplementary Figure 3**, we explain why here negative epistasis between nonessential genes seems to be more abundant than expected.

Saccharomyces cerevisiae is another species whose reconstructed high-quality metabolic networks have been extensively validated experimentally[19,21]. We repeated the above FBA in *S. cerevisiae* and obtained similar general findings on the frequencies of positive and negative epistasis and the functional relationships of epistatic reactions (**Fig. 3** and **Table 1**). Specifically, only 0.2% of reaction pairs show negative epistasis ($\tilde{\varepsilon} \leq -0.01$), 83% of which have functional overlaps. By contrast, >95% of reaction pairs show positive epistasis ($\tilde{\varepsilon} \geq 0.01$), but only 20% of those have overlapping functions.

Our computational results seem to be robust against several potential caveats to the computational analysis (**Supplementary Note**). Because of the difficulty of conducting partial gene deletion in *E. coli*, we next pursued experimental validation of our computational predications in *S. cerevisiae*, examining six essential and two nonessential genes from seven functional categories (**Supplementary Tables 2** and **3**). We deleted one allele per gene from a diploid *S. cerevisiae* strain to achieve partial disruption of a gene. For these experiments, we used haploinsufficient genes to ensure that partial gene disruption would affect fitness. As such, we examined only nonmetabolic genes, because metabolic genes are rarely haploinsufficient[26]. Nonmetabolic genes are expected to behave similarly to metabolic genes in terms of epistasis[27], as long as the final product is composed of multiple constituents with a fixed or preferred composition stoichiometry. We measured the fitness of each strain through a growth competition assay with a reference strain, followed by cell counting using FACS. We then calculated the fitness values of all single-deletion strains and all pairwise double-deletion strains relative to the wild type, which allowed us to estimate the amount of epistasis between genes (see Online Methods and the **Supplementary Note**). Among the 27 gene pairs that involve at least one essential gene, 23 (85%) have significantly positive ε ($P < 0.05$, *t*-test), two have significantly negative ε, and the remaining two do not show significant epistasis (**Fig. 4a**). The mean $\tilde{\varepsilon}$ among the 23 positively epistatic pairs is 0.78, and 11 of them have $\tilde{\varepsilon}$ not significantly smaller than 1. The epistasis between the two nonessential genes is not statistically significant. These results strongly support the general findings of our computational predictions that essential genes often show epistasis with functionally unrelated genes.

Because the above experiment could not examine haplosufficient genes, we used the newly developed decreased abundance by mRNA perturbation (DAmP) method[28] to mimic partial gene deletion, in which a marker gene is inserted into the 3′ untranslated region of a gene such that its protein expression may be reduced to <50%. We studied nine haplosufficient genes belonging to eight functional categories, including four essential genes that are knocked down by DAmP and five nonessential genes that are knocked out (**Supplementary Table 2**). We were able to measure the epistasis of 33 of the 36 gene pairs in haploid cells (**Fig. 4b**, **Supplementary Table 4** and **Supplementary Fig. 3**). Of the 23 gene pairs that have epistasis estimates and involve at least one essential gene, 20 (87%) showed significantly positive ε ($P < 0.05$, *t*-test) and two showed significantly negative ε; the remaining one did not show significant epistasis (**Fig. 4b**). These results further support our computational result of abundant positive epistasis involving essential genes, even among functionally unrelated ones. In the **Supplementary Note**, we discuss possible explanations for why selected previous studies examining the extent of epistasis in *E. coli*, yeast and other species did not find a comparably high prevalence of positive epistasis[1–3,15,17,29].

In summary, our flux balance analysis of the *E. coli* and yeast metabolic networks and the subsequent experimental validations for 61 gene pairs in *S. cerevisiae* reveals a high prevalence of positive epistasis involving essential genes. Whereas negative epistasis was usually found among genes involved in reactions with overlapping functions, positive epistasis often occurred among genes involved in reactions with unrelated functions. The proportion of essential genes is ~7% in *E. coli*, 17% in *S. cerevisiae* and 55% in mouse[30], and positive epistasis is therefore likely to be even more prevalent in higher eukaryotes than we found here. These findings suggest the distinction of genetic interaction from nonmultiplicative (or nonadditive) gene effects and provide a note of caution against the use of positive epistasis to infer genetic pathways and gene-gene interactions. Although one may argue that, because all metabolic genes share functions in supporting cell growth, their epistasis is not surprising, we suggest that, if epistasis corresponds to such a crude functional relationship, it provides little biological insight. Although our results are presented primarily using $\tilde{\varepsilon}$, it is clear that positive epistasis is highly abundant and much more prevalent than negative epistasis, even when ε is used (**Supplementary Figs. 4** and **5**). This is also the case when the majority of mutations are only slightly deleterious (**Supplementary Table 5**). These observations also suggest the need for re-evaluation of evolutionary theories that depend on overall negative epistasis, such as the mutational deterministic hypothesis of the evolution of sexual reproduction[11] and the hypothesis of reduction in mutational load by truncation selection against deleterious mutations[14].

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank N. Bharucha, G. Kalay, A. Kumar, J. Ma and B. Williams for advice and assistance in yeast experiments, B. Palsson and his group for instruction on FBA, B.-Y. Liao for drawing **Supplementary Figure 2** and M. Bakewell, S. Cho, W. Grus, B.-Y. Liao, and C. Maclean for valuable comments. This work was supported by research grants from the US National Institutes of Health and University of Michigan Center for Computational Medicine and Biology to J.Z.

AUTHOR CONTRIBUTIONS

X.H. and J.Z. conceived the research; X.H., Z.W., W.Q. and J.Z. designed the experiments; X.H., W.Q., Z.W., Y.L. and J.Z. conducted the experiments; X.H., W.Q., Z.W. and J.Z. analyzed the data; X.H. and J.Z. drafted the manuscript and all authors contributed to the final manuscript writing and its revisions.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Tischler, J., Lehner, B. & Fraser, A.G. Evolutionary plasticity of genetic interaction networks. *Nat. Genet.* **40**, 390–391 (2008).
2. Boone, C., Bussey, H. & Andrews, B.J. Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449 (2007).
3. Roguev, A. *et al.* Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* **322**, 405–410 (2008).
4. Moore, J.H. & Williams, S.M. Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *Bioessays* **27**, 637–646 (2005).
5. Cordell, H.J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.* **11**, 2463–2468 (2002).
6. Phillips, P.C. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867 (2008).
7. Bateson, W. *Mendel's Principles of Heredity* (Cambridge University Press, Cambridge, UK, 1909).
8. Fisher, R.A. The correlations between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**, 399–433 (1918).

9. Hartman, J.L., Garvik, B. & Hartwell, L. Principles for the buffering of genetic variation. *Science* **291**, 1001–1004 (2001).
10. Coyne, J.A. Genetics and speciation. *Nature* **355**, 511–515 (1992).
11. Kondrashov, A.S. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440 (1988).
12. Barton, N.H. & Charlesworth, B. Why sex and recombination? *Science* **281**, 1986–1990 (1998).
13. Kondrashov, A.S. & Crow, J.F. Haploidy or diploidy: which is better? *Nature* **351**, 314–315 (1991).
14. Crow, J.F. & Kimura, M. Efficiency of truncation selection. *Proc. Natl. Acad. Sci. USA* **76**, 396–399 (1979).
15. Jasnos, L. & Korona, R. Epistatic buffering of fitness loss in yeast double deletion strains. *Nat. Genet.* **39**, 550–554 (2007).
16. Yeh, P., Tschumi, A.I. & Kishony, R. Functional classification of drugs by properties of their pairwise interactions. *Nat. Genet.* **38**, 489–494 (2006).
17. Segrè, D., Deluna, A., Church, G.M. & Kishony, R. Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83 (2005).
18. Price, N.D., Reed, J.L. & Palsson, B.O. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2**, 886–897 (2004).
19. Papp, B., Pal, C. & Hurst, L.D. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**, 661–664 (2004).
20. Ibarra, R.U., Edwards, J.S. & Palsson, B.O. *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature* **420**, 186–189 (2002).
21. Harrison, R., Papp, B., Pal, C., Oliver, S.G. & Delneri, D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc. Natl. Acad. Sci. USA* **104**, 2307–2312 (2007).
22. Deutscher, D., Meilijson, I., Kupiec, M. & Ruppin, E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat. Genet.* **38**, 993–998 (2006).
23. Edwards, J.S., Ibarra, R.U. & Palsson, B.O. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* **19**, 125–130 (2001).
24. Reed, J.L., Vo, T.D., Schilling, C.H. & Palsson, B.O. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).
25. Szathmáry, E. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. *Genetics* **133**, 127–132 (1993).
26. Deutschbauer, A.M. *et al.* Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925 (2005).
27. Kishony, R. & Leibler, S. Environmental stresses can alleviate the average deleterious effect of mutations. *J. Biol.* **2**, 14 (2003).
28. Breslow, D.K. *et al.* A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat. Methods* **5**, 711–718 (2008).
29. Elena, S.F. & Lenski, R.E. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395–398 (1997).
30. Liao, B.Y. & Zhang, J. Mouse duplicate genes are as essential as singletons. *Trends Genet.* **23**, 378–381 (2007).

ONLINE METHODS

Flux balance analysis and minimization of metabolic adjustment (MOMA). Details of FBA and MOMA were as described[18,31]. We used the optimization package CPLEX on the MATLAB platform to solve the programming problems. We used the *E. coli* metabolic model of iJR904 (ref. 24) and *S. cerevisiae* model of iMM904 (ref. 32).

Measuring epistasis and identifying the functions of each reaction. We first describe the analysis in *E. coli*. To delete reaction *i*, we set both of its upper-bound and lower-bound flux constraints to zero in FBA. To perturb a reaction, we set its upper-bound flux as 50% of its wild-type optimal flux. Essential reactions constitute 212 out of 270 (78%) of all important reactions. We also studied the effect of different degrees of perturbation. Of the 270 reactions, 11 have variable optimal fluxes in the wild type. For these reactions, we used the minimal optimal fluxes so that any constraint in flux would be deleterious, allowing us to measure epistatic effects of deleterious mutations. Note that constraining the flux of a reaction to ≤50% of its wild-type level is not equivalent to constraining the enzyme concentration to ≤50% of its wild-type level, owing to the nonlinear relationship between enzyme concentration and flux[33].

We conducted all pairwise double perturbations of the 270 important reactions *in silico*. The relative fitness of a mutant is defined as the maximal biomass production rate of the mutant, relative to that of the wild type[17]. Epistasis is measured by the equation $\varepsilon = W_{XY} - W_X W_Y$, where W_X and W_Y represent the fitness values of two single mutants relative to the wild type, and W_{XY} represents the fitness of the corresponding double mutant[17]. Scaled epistasis[17] is defined by $\tilde{\varepsilon} = W_{XY}/(W_X W_Y) - 1$ when $\varepsilon < 0$ and $\tilde{\varepsilon} = (W_{XY} - W_X W_Y)/[\min(W_X, W_Y) - W_X W_Y]$ when $\varepsilon > 0$. Thus, $\tilde{\varepsilon}$ is normally between −1 and 1, although it can be >1 if W_{XY} is greater than $\min(W_X, W_Y)$.

When the fluxes of two reactions are constrained simultaneously, if both reactions have variable optimal fluxes, it is possible that their minimal optimal fluxes cannot be simultaneously realized in the wild type. In such instances, we actually overconstrained one of the two reactions in measuring W_{XY}, rendering ε as an underestimate so that our conclusion of prevalent positive epistasis was conservative. However, among the 255 reactions presented in **Figure 2**, only four had variable optimal fluxes, and the pairwise ε values among them were all non-negative. Note that our epistasis measurement is completely independent from the identification of the function of each reaction. In other words, the observed abundance of positive epistasis is not dependent on the assumptions made in identifying the functions of metabolic reactions. The function of each reaction was identified by a removal-recovery method (see the **Supplementary Note** for details).

Measuring epistasis and identifying functions of yeast metabolic reactions. The yeast metabolic network contains 1,412 biochemical reactions, including 538 dead-end reactions. Using FBA, we found that 244 reactions had fitness effects upon deletion in cells grown in glucose minimal medium. We were able to identify the functions for 212 (158 essential and 54 nonessential) of these 244 reactions unambiguously, using the removal-recovery method. We measured epistasis between reactions by the same method used for the *E. coli* network.

Yeast strains and growth conditions. In this study we used haploid *S. cerevisiae* strain BY4742 (*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*) and diploid strain BY4743 (*MAT***a**/*MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/+ met15Δ0/+ ura3Δ0/ura3Δ0*), both derived from the laboratory strain s288c[34]. The strains were grown on rich YPD medium (1% yeast extract, 2% peptone, 2% glucose) or minimal synthetic dextrose (SD) medium (0.67% yeast nitrogen base without amino acids, 2% glucose) with appropriate dropout supplements (Clontech). We made 5-fluoroorotic acid (5-FOA) agar medium by mixing −Ura dropout supplement (Clontech), uracil (Sigma; final concentration 50 mg l^{-1}) and 5-FOA (Research Products International; final concentration 0.1% w/v) into SD agar media.

Strain constructions. We used two strategies. In the first, we used the diploid strain BY4743 (ref. 34) as the wild-type strain and used either *URA3* or *LEU2* to replace one allele of a chosen gene in BY4743. Because the YPD medium supplies both uracil and leucine, the replacement of target genes with a functional *URA3* or *LEU2* gene is expected to have a minimal fitness effect (see below for details). Note that the *LEU2* marker used in this study excludes the flanking tRNA genes that are commonly included in the *LEU2* marker. We amplified *URA3* from pRS416 (Stratagene) and *LEU2* from PRS305 (American Type Culture Collection) using gene-specific primers containing ~60 nucleotides matching the sequences upstream and downstream of the open reading frame (ORF) of the gene to be replaced. We used the resulting cassette to replace the target gene using a homologous recombination–based method[35]. We selected *URA3*-inserted strains on a uracil-dropout synthetic media (SD –Ura), and we selected *LEU2*-inserted strains on a leucine-dropout synthetic medium (SD –Leu). Each target gene was independently replaced with both *LEU2* and *URA3*. Thus, we made 16 single-gene replacement strains. We confirmed the status of heterozygous replacement by PCR. Because six of the eight genes are essential for growth in YPD, mating-based methods[36] cannot be used to make double-replacement strains. Instead, we made all 28 double-replacement strains by sequential replacement of two target genes with the two marker genes in diploid cells.

In the second strategy, we made DAmP strains following the original design[28], except that we used *URA3* or *LEU2* as markers rather than the kanamycin-resistance (*Kan*^{R}) cassette. We inserted the marker gene exactly after the stop codon of each gene. For nonessential genes, we deleted the ORF (from start codon to stop codon) by either *URA3* or *LEU2*. We used haploid BY4742 as the wild-type strain in this approach.

The reference strain was marked with the Venus variant of yellow fluorescent protein (vYFP)[37] for FACS. We amplified the vYFP sequence from pBS7 (Yeast Resource Center, Univ. Washington) and introduced it into plasmid p426GPD[38] using the EcoRI and BamHI sites. vYFP proteins are expressed from an extremely strong promoter, *GPD*, in yeast (the promoter of *TDH3*) and with the *CYC1* terminator. We first replaced the ORF of *MET15* in BY4742 with *URA3* by a PCR-based gene-replacement method[35] and selected it on SD –Ura. We then replaced *URA3* with the vYFP gene (together with the *GPD* promoter and *CYC1* terminator) and selected for it on 5-FOA plates. Yellow fluorescence was confirmed by live-cell fluorescence microscopy. All gene replacement strains were confirmed by genomic DNA extraction and PCR.

Growth assay and experimental estimation of epistasis. We used a growth competition assay to measure the fitness of each strain and then estimated epistasis and its confidence interval (see **Supplementary Note** for details).

Comparison with previous fitness data of heterozygous deletion strains. The fitness (*W*) values of the single heterozygous deletion strains estimated from our competition assay differ from previous estimates[26,39]. For two reasons, our results are more reliable than previous results. First, we measured the fitness based on counting hundreds of thousands of cells by flow cytometry, a method that is much more reliable than the microarray-based method used in the previous fitness estimation. Second, we compared the deletion strains to a wild-type strain, a step that was not included in previous fitness estimations.

No epistasis between the selectable markers for gene replacement. The high prevalence of positive epistasis observed in the experiment could potentially be a result of epistasis between the two selectable markers used in gene replacement (*LEU2* and *URA3*). To exclude this possibility, we used *LEU2* and *URA3* to each replace one allele of a nonfunctional gene (*HO*) in the diploid strain BY4743 and measured the epistasis between *LEU2* and *URA3*. *HO* encodes a site-specific endonuclease required for gene conversion at the *MAT* locus (homothallic switching). We selected *HO* for two reasons. First, if we simply inserted the two marker genes into an intergenic region, it is possible that the marker genes could destroy unknown functional elements in the region[40] and cause unwanted fitness effects. Second, the *HO* gene in BY4743 and its ancestor s288c has several severe mutations and is apparently nonfunctional[41,42]. Thus, replacement with our marker genes would not have any unwanted side effect. We found no significant epistasis between *LEU2* and *URA3*. The fitness of the *URA3* insertion strain is $W_{URA3} = 1.014$, that of the *LEU2* insertion strain is $W_{LEU2} = 1.003$ and that of the *URA3* and *LEU2* double-insertion strain is $W_{URA3\text{-}LEU2} = 1.016$. The epistasis between *URA3* and *LEU2* is $\varepsilon = -0.001$ ($P > 0.9$, *U*-test). Furthermore, the absolute value of ε between the two marker genes is small compared with most of the epistasis values observed (**Supplementary Table 3**).

URL. Additional analyses related to this publication can be found at http://www.umich.edu/~zhanglab/download.htm.

31. Segrè, D., Vitkup, D. & Church, G.M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. USA* **99**, 15112–15117 (2002).
32. Mo, M.L., Palsson, B.O. & Herrgard, M.J. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Syst. Biol.* **3**, 37 (2009).
33. Kacser, H. & Burns, J.A. The molecular basis of dominance. *Genetics* **97**, 639–666 (1981).
34. Brachmann, C.B. *et al.* Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115–132 (1998).
35. Amberg, D.C., Burke, D.J. & Strathern, J.N. *Methods in Yeast Genetics. A Cold Spring Harbor Laboratory Course Manual* 59–76 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 2005).
36. Tong, A.H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368 (2001).
37. Nagai, T. *et al.* A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87–90 (2002).
38. Mumberg, D., Muller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119–122 (1995).
39. Steinmetz, L.M. *et al.* Systematic screen for human disease genes in yeast. *Nat. Genet.* **31**, 400–404 (2002).
40. Nagalakshmi, U. *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349 (2008).
41. Ekino, K., Kwon, I., Goto, M., Yoshino, S. & Furukawa, K. Functional analysis of *HO* gene in delayed homothallism in *Saccharomyces cerevisiae* wy2. *Yeast* **15**, 451–458 (1999).
42. Meiron, H., Nahon, E. & Raveh, D. Identification of the heterothallic mutation in HO-endonuclease of *S. cerevisiae* using *HO*/*ho* chimeric genes. *Curr. Genet.* **28**, 367–373 (1995).

doi:10.1038/ng.524