Supporting Information

Wang et al. 10.1073/pnas.1004666107

SI Methods

Pleiotropy Datasets. The yeast morphological pleiotropy dataset (1) includes the phenotypic information acquired by fluorescent imaging of 4,718 yeast nonessential gene deletion haploid strains as well as the wild-type haploid strain. The phenotypes include 501 quantitative traits of yeast cellular morphology such as cell shape, actin cytoskeleton, and nuclear morphology. For each trait, the average phenotypic value of 200 wild-type cells was obtained from each of 126 independent cultures to estimate the wild-type mean and variance, and the average value of 200 cells was obtained from one culture for each deletion strain to estimate the phenotype of the deletion strain. The raw data were obtained from http://scmd.gi.k.u-tokyo.ac.jp/datamine/. Following the suggestion of the original authors (1), we transformed the raw data of the 501 traits by power transformation (2) and then checked for normality in distribution among wild-type cells using the Shapiro–Wilk test. For 222 traits, the phenotypes of the wild-type cells either are not power transformable or do not follow normal distributions. These traits were thus excluded from subsequent analysis and the remaining 279 traits were considered in our morphological pleiotropy data.

The size of the phenotypic effect of a gene on a trait is measured by the statistical Z-score, which is defined by $Z = (m_d - m_{wt})/SD$, where m_{wt} and SD are the mean and standard deviation of the transformed measures of the trait from wild-type cells, respectively, and m_d is the transformed measure of the trait from a cell deficient of the gene. Note that because m_d can be larger or smaller than m_{wt}, Z can be positive or negative. A given $m_d - m_{wt}$ value indicates a greater fitness effect when it occurs in a more important trait than in a less important trait. Because the SD of a trait is expected to be negatively correlated with the strength of stabilizing selection on the trait (i.e., the importance of a trait to organismal fitness), Z-scores effectively standardize phenotypic effects in terms of fitness effects and thus are comparable among traits. To determine the number of traits a gene affects, we calculated the statistical P values according to the Z-scores using the standard normal distribution. Because we simultaneously tested 279 traits for each gene, we corrected for multiple testing using a 5% false discovery rate (FDR). In other words, if a trait shows a Q value <5% in a gene-deletion strain, we consider that this trait is affected by this gene. By this cutoff, a gene affects on average 22 traits. Thus, the number of false positives is normally <1 trait per gene. We also used the more conservative Bonferroni correction of multiple testing, and the results are shown in Fig. S3. After the removal of genes that do not affect any trait and traits that are not affected by any gene, the yeast morphological dataset contains 2,449 genes and 253 traits. The information on the fitness effects of individual gene deletions in yeast was obtained from http://www-deletion.stanford.edu/YDPM/YDPM_index.html.

The same collection of yeast gene deletion strains was also screened under 22 different environmental conditions for growth defects (3). A gene is considered to affect growth under a condition when the deletion strain shows significantly slower growth than the wild-type strain. Because the data did not contain quantitative measures of growth rates, the gene–trait relationship is qualitative. That is, a gene either affects or does not affect a trait. In total, 774 genes affect growth in at least 1 of the 22 environmental conditions. This dataset is referred to as the yeast environmental pleiotropy dataset.

We also obtained yeast knockout phenotype information from the Comprehensive Yeast Genome Database (CYGD) (4), which catalogs literature-curated physiological defects of yeast gene deletion strains from small-scale experiments. After removing phenotypes that are annotated as "unclassified," we obtained our yeast physiological pleiotropy data containing 1,256 genes that affect ≥1 of 120 traits. As in the yeast environmental pleiotropy dataset, this dataset only has qualitative information about gene–trait relationships.

To identify genes required for early embryogenesis in nematodes, a recent study used genome-wide RNA-mediated interference (RNAi) to silence gene expression in early *C. elegans* embryos (5). The targeted RNAi experiment for each gene was repeated in six embryos, and 45 phenotypic traits were screened for developmental defects. We consider that a gene affects a trait if at least two of the six embryos showed phenotypic defects. After the removal of one trait named "complex phenotype," we obtained our nematode pleiotropy dataset including 661 genes that affect ≥1 of 44 traits. This dataset provides only qualitative information about gene–trait relationships.

The mouse pleiotropy data were derived from annotations of MGI version 4.2 (http://www.informatics.jax.org/) (6). At the time of this study, 5,586 mouse genes were annotated with one or more Mammalian Phenotype (MP) IDs indicating the phenotypes when the genes were knocked out, knocked down, mutated by transgenic insertions, or occasionally mutated by point mutations. MP IDs are hierarchically structured. That is, one parent MP ID (e.g., MP:0002102, abnormal ear morphology) represents a phenotype lineage that may include several child MP IDs to describe a more detailed phenotype (e.g., MP:0000026, abnormal inner ear morphology; MP:0002177, abnormal outer ear morphology). Here, we used 308 parent MP IDs to define the pleiotropy of mouse genes. These 308 MP IDs were manually selected using the criterion that each MP ID should be phenotypically distinct, if not independent, from the other MP IDs. If a mouse gene is annotated for a child MP, its parent MP ID is used. Consequently, pleiotropy of 4,915 mouse genes associated with at least 1 of the 308 MP IDs was obtained. This dataset provides only qualitative information of the gene–trait relationships.

Simulating Normally Distributed Phenotypic Effects of Genes. For a given gene i, we first calculated the SD (σ_i) of its phenotypic effect size distribution from the yeast morphological pleiotropy data. Note that, in this calculation, we used the phenotypic effects of the gene on all 279 traits, regardless of whether these effects are statistically significant. We then randomly generated this gene's phenotypic effects on each of the 279 traits using a normal distribution with mean = 0 and SD = σ_i. We performed these steps for all 4,718 genes to produce a 4,718 × 279 random effect-size matrix. We then analyzed this simulated dataset following the analysis of the real data.

To examine the impact of different SDs of different genes on our results, we conducted the second simulation. The procedure is the same as the above simulation, except that, instead of using different SDs for different genes, we used the same SD for all genes. This SD used was the mean SD of all genes in the actual data.

Principal Component Analysis of Phenotypic Traits. To obtain independent phenotypic traits, we performed a principal component analysis of the wild-type phenotypic matrix from the yeast morphological pleiotropy data. The wild-type phenotypic matrix W, in which the 126 rows represent independent wild-type cell cultures and the 279 columns correspond to the original phenotypic traits, was first standardized into Z-scores by each column. The principal component analysis was conducted using MATLAB to obtain the principal

component coefficient matrix C. After the linear transformation with the coefficient matrix, the compound phenotypic matrix $W' = W \times C$ becomes orthogonal. We then applied this linear transformation to the standardized mutant phenotypic matrix M to obtain the compound mutant phenotypic matrix $M' = M \times C$. In M and M', each row is a deletion strain and each column is a trait.

Genes Affecting More Traits Have Larger Per-Trait Effects on Average. Comparing two genes both with normal distributions of effect sizes but with different SDs, here we prove mathematically that the gene with the larger SD affects more traits (when an effect-size cutoff is applied) and has on average a larger per-trait effect.

For a given gene, let $f(x)$ be the probability density function of the distribution of effect size, where effect size is measured by Z-scores. On the basis of empirical observations, we assume that $f(x)$ is a normal distribution with mean = 0 and SD = t, or

$$f(x) = \frac{1}{\sqrt{2\pi}t} e^{-x^2/(2t^2)}. \quad \text{[S1]}$$

Let $g > 0$ be the cutoff used to determine whether a trait is regarded as being affected significantly by the gene. The mean effect size per trait $F(t)$ can be expressed as

$$F(t) = \frac{\int_g^{+\infty} xf(x)dx}{\int_g^{+\infty} f(x)dx} = \frac{u(t)}{v(t)}, \quad \text{[S2]}$$

where $u(t) = \int_g^{+\infty} xf(x)\,dx$ and $v(t) = \int_g^{+\infty} f(x)dx$. Below, we prove that $F(t)$ is a monotonically increasing function of t, or $F'(t) > 0$.

We have

$$F'(t) = \frac{u'(t)v(t) - u(t)v'(t)}{v^2(t)} = \frac{v(t) - u(t)v'(t)/u'(t)}{v^2(t)} = \frac{A}{B}, \quad \text{[S3]}$$

where $A = v(t) - u(t)v'(t)/u'(t)$ and $B = v^2(t)$. We can derive that

$$\begin{aligned}
u(t) &= \int_g^{+\infty} \frac{x}{\sqrt{2\pi}t} e^{\frac{-x^2}{2t^2}} dx = \frac{t}{\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}};\\
u'(t) &= \frac{1}{\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}} \left(1 + \frac{g^2}{t^2}\right);\\
v(t) &= \int_g^{+\infty} \frac{1}{\sqrt{2\pi}t} e^{\frac{-x^2}{2t^2}} dx = \frac{1}{\sqrt{2\pi}} \int_g^{+\infty} \frac{1}{t} e^{\frac{-x^2}{2t^2}} dx\\
&= \frac{1}{\sqrt{2\pi}} \int_{\frac{g}{t}}^{+\infty} e^{\frac{-y^2}{2}} dy, \text{ where } y = x/t;\\
v'(t) &= \frac{g}{t^2\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}}.
\end{aligned} \quad \text{[S4]}$$

Therefore,

$$\begin{aligned}
A &= \frac{1}{\sqrt{2\pi}} \int_{\frac{g}{t}}^{+\infty} e^{\frac{-y^2}{2}} dy - \left(\frac{t}{\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}}\right)\\
&\quad \times \left(\frac{a}{t^2\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}}\right) \Big/ \left[\frac{1}{\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}} \left(1 + \frac{g^2}{t^2}\right)\right]\\
&= \frac{1}{\sqrt{2\pi}} \int_{\frac{g}{t}}^{+\infty} e^{\frac{-y^2}{2}} dy - \frac{gt}{(g^2 + t^2)\sqrt{2\pi}} e^{\frac{-g^2}{2t^2}}\\
&= \frac{1}{\sqrt{2\pi}} \left(\int_h^{+\infty} e^{\frac{-y^2}{2}} dy - \frac{h}{1+h^2} e^{-h^2/2}\right), \text{where } h = g/t.
\end{aligned} \quad \text{[S5]}$$

Because it can be shown that

$$\begin{aligned}
\int_h^{+\infty} e^{-y^2/2} dy &= \frac{h^2}{1+h^2} \int_h^{+\infty} \left(1 + \frac{1}{h^2}\right) e^{-y^2/2}\, dy\\
&> \frac{h^2}{1+h^2} \int_h^{+\infty} \left(1 + \frac{1}{y^2}\right) e^{-y^2/2}\, dy\\
&= -\frac{h^2}{1+h^2} \int_h^{+\infty} d\left(\frac{1}{y} e^{-y^2/2}\right)\\
&= \frac{h}{1+h^2} e^{-\frac{h^2}{2}},
\end{aligned} \quad \text{[S6]}$$

A is positive. Because B is also positive, $F'(t)$ is positive. In other words, $F(t)$ is a monotonically increasing function of t.

Let N be the total number of traits considered. Then the number of traits affected by a gene is $n(t) = Nv(t)$. Because $v'(t)$ is positive, n is a monotonically increasing function of t. Thus, both $F(t)$ and $n(t)$ increase with t. In other words, when t is larger, both the number of affected traits and the mean effect size increase, which creates the phenomenon of larger per-trait effect sizes for genes affecting more traits. Although in the above proof only traits with Z-scores larger than a positive cutoff g are considered to be affected by a gene, the result is the same when traits with Z-scores smaller than a negative cutoff g are considered to be affected, because $f(x)$ is symmetrical to 0. Thus, when all traits with absolute Z-scores larger than a cutoff $g > 0$ are considered to be affected, which is what we did in actual data analysis, the above proof is also valid.

Note that our proof assumes that we use a constant cutoff $g > 0$ for all genes. In the actual data analysis, the cutoff may vary for different genes when the same false discovery rate is used to determine the cutoff. However, the small variation in cutoff apparently did not affect the general trend of larger per-trait effect sizes for genes affecting more traits.

Also note that here we used a normal distribution to model the effect sizes of a gene on various traits because the normality is what we empirically observed (Fig. 3*A*). The normality is not necessary for the phenomenon that genes affecting more traits have on average larger per-trait effects.

Existence of Nonzero $n_{optimal}$ When $b > 0.5$. Let T_E be the total phenotypic effect size of a mutation measured by the Euclidian distance and n be the degree of pleiotropy (or effective organismal complexity). Here we prove that when the exponent $b > 0.5$ in the scaling relationship of $T_E = an^b$, the highest adaptation rate occurs at an intermediate n. On the basis of Orr (7), the adaptation rate of a population is

$$U(n) = \frac{dw}{dt} = -\frac{4kT_E^2}{n} M w \ln w = -4ka^2 n^{2b-1} M w \ln w, \quad \text{[S7]}$$

where k is a positive constant dependent on population size and mutation rate, $0 < w < 1$ is the current mean fitness of the population, $M = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} (y-x)^2 e^{\frac{-y^2}{2}} dy$, and $x = \frac{T_E\sqrt{n}}{2\sqrt{-2\ln w}} = \frac{an^{b+0.5}}{2\sqrt{-2\ln w}}$. We can show that

$$\begin{aligned}
U'(n) &= \left(\frac{-4ka^2 w \ln w}{\sqrt{2\pi}}\right) \frac{d\left[n^{2b-1}(\sqrt{2\pi}M)\right]}{dn}\\
&= \left(\frac{-4ka^2 w \ln w}{\sqrt{2\pi}}\right) \left[(2b-1)n^{2b-2}(\sqrt{2\pi}M) + n^{2b-1} \frac{d(\sqrt{2\pi}M)}{dx} \frac{dx}{dn}\right]\\
&= \left(\frac{-4ka^2 n^{2b-2} w \ln w}{\sqrt{2\pi}}\right) \left[(2b-1)(\sqrt{2\pi}M) + n \frac{d(\sqrt{2\pi}M)}{dx} \frac{x}{n} \left(b + \frac{1}{2}\right)\right]\\
&= \left(\frac{-4ka^2 n^{2b-2} w \ln w}{\sqrt{2\pi}}\right) \left[(2b-1)(\sqrt{2\pi}M) + x\left(b + \frac{1}{2}\right) \frac{d(\sqrt{2\pi}M)}{dx}\right].
\end{aligned} \quad \text{[S8]}$$

It can be shown by Maxima (http://maxima.sourceforge.net/), a computer algebra system, that

$$\frac{d(\sqrt{2\pi}M)}{dx} = \frac{d\left[-xe^{-x^2/2} + \sqrt{\frac{\pi}{2}}(1+x^2)Erfc(\frac{x}{\sqrt{2}})\right]}{dx} = -2e^{-x^2/2} + \sqrt{2\pi}xErfc(\frac{x}{\sqrt{2}}), \quad \textbf{[S9]}$$

where Erfc(x) is the complementary error function:

$$Erfc(x) = 1 - Erf(x) = \frac{2}{\sqrt{\pi}}\int_x^{+\infty} e^{-t^2}dt. \quad \textbf{[S10]}$$

Combining Eqs. **S8** and **S9**, we have

$$U'(n) = \frac{-4ka^2n^{2b-2}w\ln w}{\sqrt{2\pi}}m(x), \quad \textbf{[S11]}$$

$$\text{where } m(x) = \sqrt{\frac{\pi}{2}}\,Erfc\left(\frac{x}{\sqrt{2}}\right)(4bx^2 + 2b - 1) - 4bxe^{-x^2/2}. \quad \textbf{[S12]}$$

When b =0, $m(x) = -\sqrt{\frac{\pi}{2}}Erfc(\frac{x}{\sqrt{2}}) < 0$. Thus, $U'(n) < 0$. This result means that $U(n)$ decreases with n. Let n_{optimal} be the n with the largest U. Our results indicate that $n_{\text{optimal}} = 0$.

When $b = 0.5$, we can show that

$$m(x) = 2x\left(\sqrt{\frac{\pi}{2}}\,xErfc\left(\frac{x}{\sqrt{2}}\right) - e^{\frac{-x^2}{2}}\right) = 2x\left(x\int_x^{+\infty} e^{-t^2/2}dt - e^{-x^2/2}\right) = 2x\left(\int_x^{+\infty} xe^{-t^2/2}dt - \int_x^{+\infty} te^{-t^2/2}dt\right) < 0.$$

The last step is true because x is biologically meaningful only when it is positive. This result means that $U'(n) < 0$ and $U(n)$ decreases with n. In other words, $n_{\text{optimal}} = 0$.

When $b > 0.5$, we have

$$m(0) = \sqrt{\frac{\pi}{2}}\,Erfc(0)(2b-1) - 0 > 0$$

and

$$m(2) = 4b\left[\sqrt{\frac{\pi}{2}}\,Erfc\left(\frac{x}{\sqrt{2}}\right)\left(x^2 + \frac{1}{2} - \frac{1}{4b}\right) - xe^{-x^2/2}\right] < 4b\left[\sqrt{\frac{\pi}{2}}\,Erfc\left(\frac{x}{\sqrt{2}}\right)\left(x^2 + \frac{1}{2}\right) - xe^{-x^2/2}\right] = 4b\left[\sqrt{\frac{\pi}{2}}\,Erfc(\sqrt{2})\left(\frac{9}{2}\right) - 2e^{-2}\right] = -0.0564b < 0.$$

Because $m(x)$ is a continuous function, $0 < x_{\text{optimal}} < 2$ exists for which $m(x) = 0$ and $U'(n) = 0$. As x moves from 0 to 2, U' changes from positive to negative, indicating that x_{optimal} corresponds to a peak of U. The n value determined by x_{optimal} thus corresponds to a peak of U and is positive. Thus, we proved that, when $b > 0.5$, a positive n_{optimal} exists.

1. Ohya Y, et al. (2005) High-dimensional and large-scale phenotyping of yeast mutants. *Proc Natl Acad Sci USA* 102:19015–19020.
2. Yeo I-K, Johnson RA (2000) A new family of power transformations to improve normality or symmetry. *Biometrika* 87:954–959.
3. Dudley AM, Janse DM, Tanay A, Shamir R, Church GM (2005) A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol Syst Biol* 1:2005.0001.
4. Güldener U, et al. (2005) CYGD: The Comprehensive Yeast Genome Database. *Nucleic Acids Res* 33(Database issue):D364–D368.
5. Sönnichsen B, et al. (2005) Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature* 434:462–469.
6. Bult CJ, Eppig JT, Kadin JA, Richardson JE, Blake JA, Mouse Genome Database Group (2008) The Mouse Genome Database (MGD): Mouse biology and model systems. *Nucleic Acids Res* 36(Database issue):D724–D728.
7. Orr HA (2000) Adaptation and the cost of complexity. *Evolution* 54:13–20.

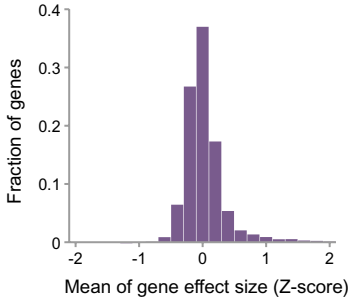

Fig. S1. Frequency distribution of the mean effect size (measured by *Z*-score) of a gene on the 279 morphological traits for all 4,718 yeast genes. Note that the effect of a gene on a trait can be either positive or negative.

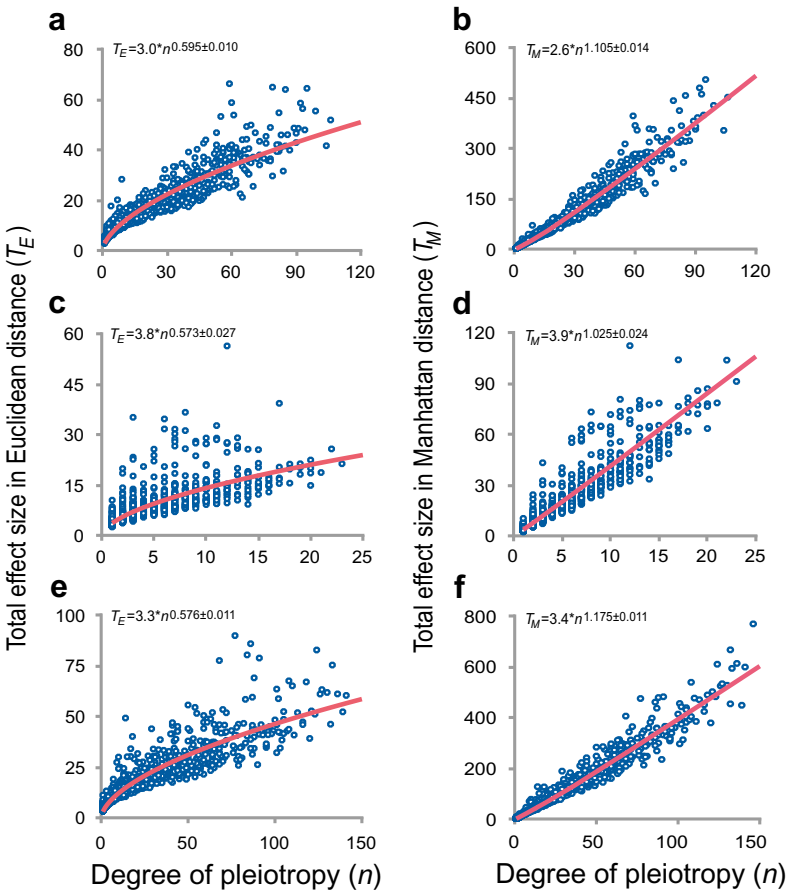

Fig. S2. The phenomenon of larger per-trait effects from genes affecting more traits is robust. Observed scaling relationships between the degree of pleiotropy and (*A*) Euclidean distance or (*B*) Manhattan distance are shown, on the basis of the yeast morphological pleiotropy data from which a random 50% of the traits are removed. The orange curve is the best fit to the power function whose estimated parameters are shown in the upper left. The numbers after ± show the 95% confidence interval for the estimated scaling exponent. *C* and *D* are similar to *A* and *B* except that the dataset used is generated after the random removal of 90% of the traits. *E* and *F* are similar to *A* and *B* except that the dataset used is generated by merging traits with a Pearson's correlation coefficient in gene effects >0.7.

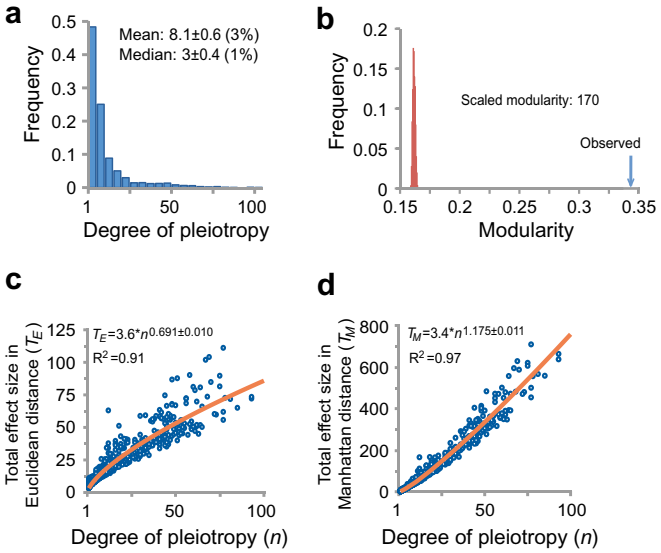

Fig. S3. Yeast morphological pleiotropy data analyzed using the conservative Bonferroni method to correct for multiple testing. (*A*) Genome-wide frequency distribution of the degree of gene pleiotropy. The numbers in parentheses are the mean and median degrees of pleiotropy divided by the total number of traits. After the removal of genes that do not affect any trait and traits that are not affected by any gene, there are 2,091 genes and 264 traits. (*B*) Observed modularity (blue arrow) and the distribution of modularity for 250 randomly rewired networks (red histograms). Observed scaling relationships between the degree of pleiotropy and the total effect size measured by (*C*) Euclidean distance or (*D*) Manhattan distance are shown. The orange curve is the best fit to the power function whose estimated parameters are shown in the upper left. The numbers after ± show the 95% confidence interval for the estimated scaling exponent. R^2 indicates the square of the correlation coefficient.

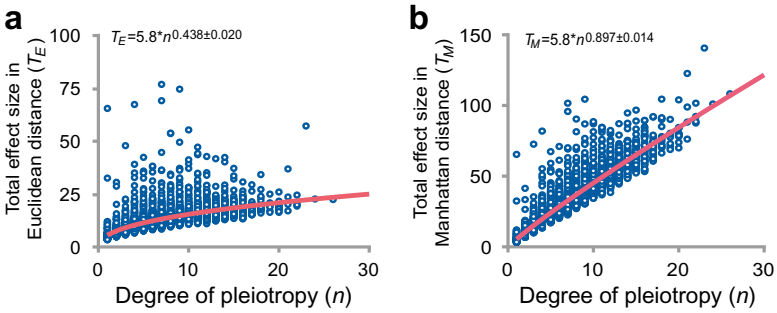

Fig. S4. Observed scaling relationships between the degree of pleiotropy and the total effect size measured by (*A*) Euclidean distance or (*B*) Manhattan distance, when the effect sizes of all genes on all traits in the actual data are randomly shuffled. The orange curve is the best fit to the power function whose estimated parameters are shown in the upper left. The numbers after ± show the 95% confidence interval for the estimated scaling exponent.

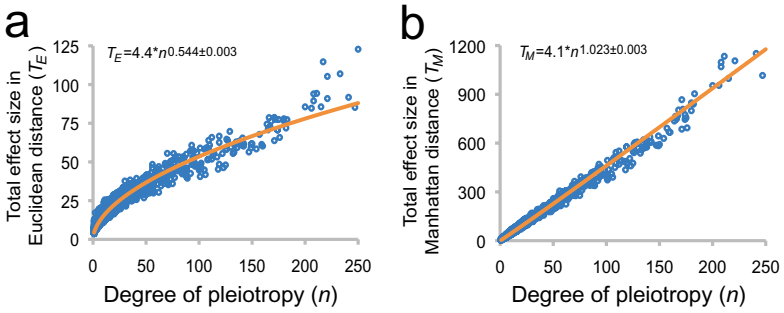

Fig. S5. Observed scaling relationships between the degree of pleiotropy and the total effect size measured by (*A*) Euclidean distance or (*B*) Manhattan distance, when the phenotypic traits are orthogonalized. The orange curve is the best fit to the power function whose estimated parameters are shown in the upper left. The numbers after ± show the 95% confidence interval for the estimated scaling exponent.

Table S1. Robustness of pleiotropy estimates

Gene pleiotropy datasets	No. of traits	Mean pleiotropy	95% confidence interval of mean pleiotropy	Mean pleiotropy divided by no. of traits	Median pleiotropy	95% confidence interval of median pleiotropy	Median pleiotropy divided by no. of traits
Yeast morphological pleiotropy							
100% of traits used	279	21.6	[20.2, 23.1]	0.077	7	[6, 8]	0.025
50% of traits used	140	10.9	[9.9, 11.9]	0.078	4	[3, 4]	0.029
10% of traits used	28	2.2	[1.7, 2.6]	0.079	1	[0, 1]	0.036
Yeast environmetnal pleiotropy							
100% of traits used	22	2.4	[1.7, 3.2]	0.109	2	[1, 3]	0.091
50% of traits used	11	1.2	[0.7, 1.8]	0.109	1	[0, 1]	0.091
10% of traits used	2	0.2	[0.1, 0.5]	0.100	0	[0, 0]	0.000
Yeast physiological pleiotropy							
100% of traits used	120	1.8	[1.3, 2.5]	0.015	1	[1, 2]	0.008
50% of traits used	60	0.9	[0.5, 1.4]	0.015	1	[0, 1]	0.017
10% of traits used	12	0.2	[0.1, 0.4]	0.017	0	[0, 0]	0.000
Nematode pleiotropy							
100% of traits used	44	4.6	[3.7, 5.6]	0.105	4	[3, 6]	0.091
50% of traits used	22	2.3	[1.7, 3.0]	0.105	2	[1, 3]	0.091
10% of traits used	4	0.5	[0.2, 0.7]	0.125	0	[0, 1]	0.000
Mouse pleiotropy							
100% of traits used	308	8.2	[7.1, 9.3]	0.027	6	[5, 6]	0.019
50% of traits used	154	4.1	[3.3, 4.9]	0.027	3	[2, 3]	0.019
10% of traits used	31	0.8	[0.5, 1.2]	0.026	1	[0, 1]	0.032

Table S2. Comparison between the observed genomic patterns of pleiotropy and assumptions made in the existing theoretical models of pleiotropy

Features	Invariant total effect model*	Euclidean superposition model†	Observed genomic patterns
Proportion of traits affected by a gene	100%	100%	1–9%
Modularity of the gene–trait network	None	None	High
Distribution of effect size on a trait	Uniform	Normal	Normal
Among-gene variation in SD of the effect-size distribution	Absent	Absent	Present
Total effect size	Constant	Increase with pleiotropy	Increase with pleiotropy
Per-trait effect size	Decrease with pleiotropy	Constant	Increase with pleiotropy
Scaling exponent *b*	0	0.5	0.6
Scaling exponent *d*	0.5	1	1.1
Degree of pleiotropy (i.e., complexity) that offers the highest adaptation rate	1	1	Intermediate level of pleiotropy

*Fisher (1) and Orr (2).
†Turelli (3), Wagner (4, 5), and Waxman and Peck (6).

1. Fisher RA (1930) *The Genetic Theory of Natural Selection* (Clarendon, Oxford), 2nd Ed.
2. Orr HA (2000) Adaptation and the cost of complexity. *Evolution* 54:13–20.
3. Turelli M (1985) Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* 111:165–195.
4. Wagner GP (1988) The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *J Evol Biol* 1:45–66.
5. Wagner GP (1989) Multivariate mutation-selection balance with constrained pleiotropic effects. *Genetics* 122:223–234.
6. Waxman D, Peck JR (1998) Pleiotropy and the preservation of perfection. *Science* 279:1210–1213.

Table S3. High modularity of gene–trait networks

Gene pleiotropy datasets	No. of traits	No. of genes	Modularity	Scaled modularity
Yeast morphological pleiotropy (all traits)	279	2,449	0.204	36.8
Yeast morphological pleiotropy (random half of the traits)	140	1,902	0.221	37.7
Yeast morphological pleiotropy (after the merge of related traits)*	197	2,272	0.209	45.0
Yeast environmetnal pleiotropy (all traits)	22	774	0.440	35.1
Yeast environmetnal pleiotropy (random half of the traits)	11	448	0.579	13.6
Yeast environmetnal pleiotropy (after the merge of related traits)*	22	774	0.440	35.1
Yeast physiological pleiotropy (all traits)	120	1,256	0.580	34.2
Yeast physiological pleiotropy (random half of the traits)	60	712	0.673	62.6
Yeast physiological pleiotropy (after the merge of related traits)*	118	1,256	0.575	27.0
Nematode pleiotropy (all traits)	44	661	0.544	50.4
Nematode pleiotropy (random half of the traits)	22	579	0.473	48.2
Nematode pleiotropy (after the merge of related traits)*	44	661	0.544	50.4
Mouse pleiotropy (all traits)	308	4,915	0.384	237.5
Mouse pleiotropy (random half of the traits)	154	4,901	0.449	197.4
Mouse pleiotropy (after the merge of related traits)*	307	4,915	0.376	202.8

*Traits with Pearson's correlation coefficient >0.7 are merged. Some datasets do not contain such correlated traits.