

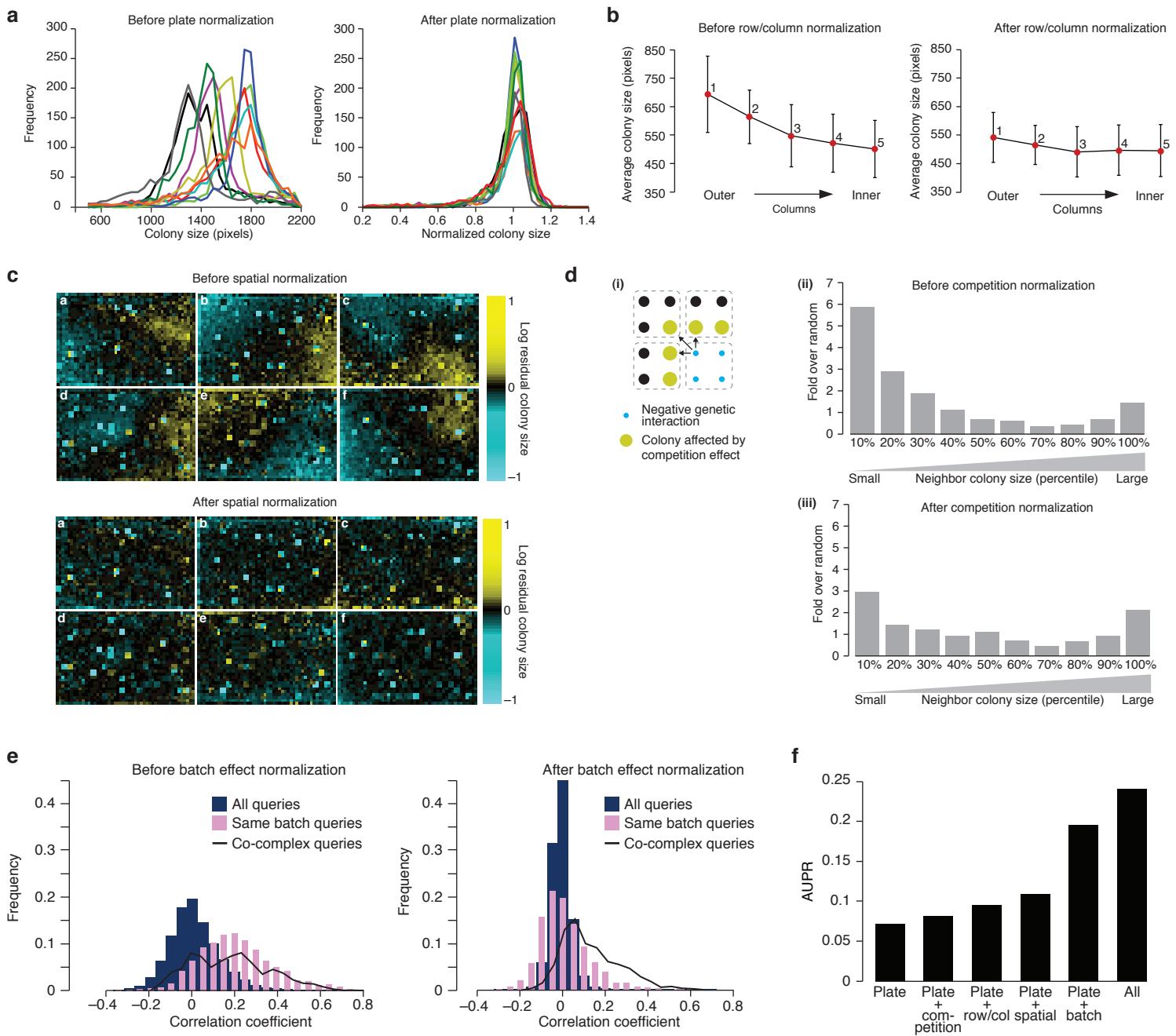
Quantitative analysis of fitness and genetic interactions in yeast on a genome scale

Anastasia Baryshnikova, Michael Costanzo, Yungil Kim, Huiming Ding, Judice Koh, Kiana Toufighi, Ji-Young Youn, Jiongwen Ou, Bryan-Joseph San Luis, Sunayan Bandyopadhyay, Matthew Hibbs, David Hess, Anne-Claude Gingras, Gary D Bader, Olga G Troyanskaya, Grant W Brown, Brenda Andrews, Charles Boone & Chad L Myers

Supplementary Figure 1	Correction of systematic experimental effects.
Supplementary Figure 2	Comparison of cross-array correlations before and after normalization.
Supplementary Figure 3	Reproducibility of reciprocal genetic interactions measured using colony sizes without normalization.
Supplementary Figure 4	Analysis of genetic interactions reported in a previous SGA study.
Supplementary Figure 5	Comparative evaluation of negative and positive genetic interactions.
Supplementary Figure 6	Batch effect contribution to the difference between SGA score and S-score functional performance.
Supplementary Figure 7	Analysis of genetic interactions within and between protein complexes.
Supplementary Figure 8	Genetic suppression between Swr1 protein complex and <i>htz1</i> .
Supplementary Figure 9	Genetic suppression between Rim101 signaling and MVB sorting pathways.
Supplementary Figure 10	Genetic suppression between FAR and TORC2 protein complexes.
Supplementary Figure 11	Genetic suppression between Polδ and COG protein complex.
Supplementary Table 1	List of normalizations.
Supplementary Table 2	List of experimental controls.
Supplementary Table 3	Sensitivity and precision of SGA genetic interaction scores.
Supplementary Note 1	SGA genetic interaction score
Supplementary Note 2	Statistical comparison of SGA genetic interaction screens to gene expression microarrays.
Supplementary Note 3	Reciprocal interaction analysis.
Supplementary Note 4	Comparison to genetic interactions curated in BioGRID.
Supplementary Note 5	Evaluating overlap of genetic and protein-protein interaction networks.
Supplementary Note 6	Protein complex suppression network.

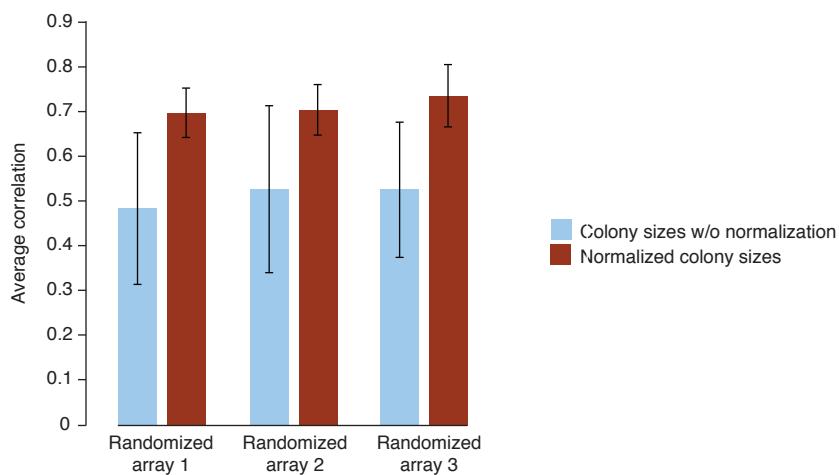
Note: Supplementary Data 1–3 and Supplementary Software are available on the Nature Methods website.

Supplementary Figure 1. Correction of systematic experimental effects



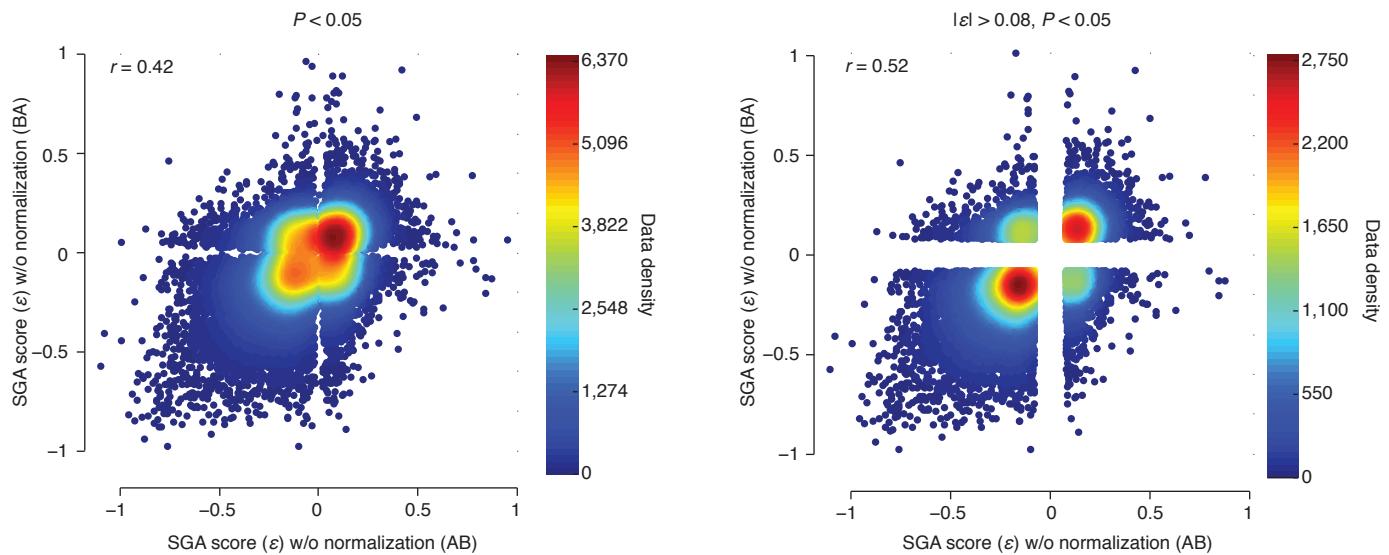
(a) Plate normalization. Colony size distributions for ten SGA plates randomly selected from control screens are shown before and after plate normalization. Each color corresponds to a different SGA plate. **(b)** Row/column normalization. Colony sizes located in columns 1 to 5 were averaged across 80 SGA control screens. The average colony size and standard deviation for each column are shown before and after row/column normalization. **(c)** Spatial effect normalization. Relative colony size was determined by measuring deviation of individual colonies from the median size for the same colony measured across 1,712 different experiments¹. Colonies that appear larger (yellow) or smaller (blue) are indicated and displayed according to plate position. Colonies that do not deviate significantly from their average size are shown in black. **(d)** Competition normalization. (i) A schematic of colonies neighboring a synthetic sick/lethal interaction (blue) is shown. Replicate colonies corresponding to the same double mutant are indicated by a dotted line. Arrows point to the colonies (yellow) that are most affected by a sick neighbor. (ii-iii) Sizes of colonies neighboring the top 1,000 largest colonies in our dataset were binned into 10 deciles and their distribution was normalized to the background distribution of all colonies in the dataset. The bar graphs show the fold enrichment over background for neighbor colonies in each decile before (ii) and after (iii) competition effect normalization. This analysis demonstrates that the largest colonies identified prior to competition correction tend to have small neighbors. The prevalence of small colonies among the neighbors of large colonies is significantly less pronounced after competition normalization. **(e)** Batch effect normalization. The distribution of Pearson correlation coefficients among genetic interaction profiles of all “query” genes (blue) and “query” genes screened within the same batch (pink) are shown before and after batch effect normalization. For reference, the distribution of correlations among genes encoding members of the same protein complex is shown (black line). Before batch correction, the average correlation of query mutants screened in the same batch is as high as the correlation of co-complex query genes. **(f)** The contribution of each normalization procedure alone was assessed by evaluating SGA scores against co-annotated genes pairs and measuring the area under the precision-recall (AUPR) curve (see Fig. 3 and Supplementary Fig. 5 for example). The complete genetic interaction dataset was scored five times by omitting all correction steps except for plate normalization and the correction procedure indicated on the x-axis. Consistent with our other evaluations, each correction alone contributes to improve data quality.

Supplementary Figure 2. Comparison of cross-array correlations before and after normalization



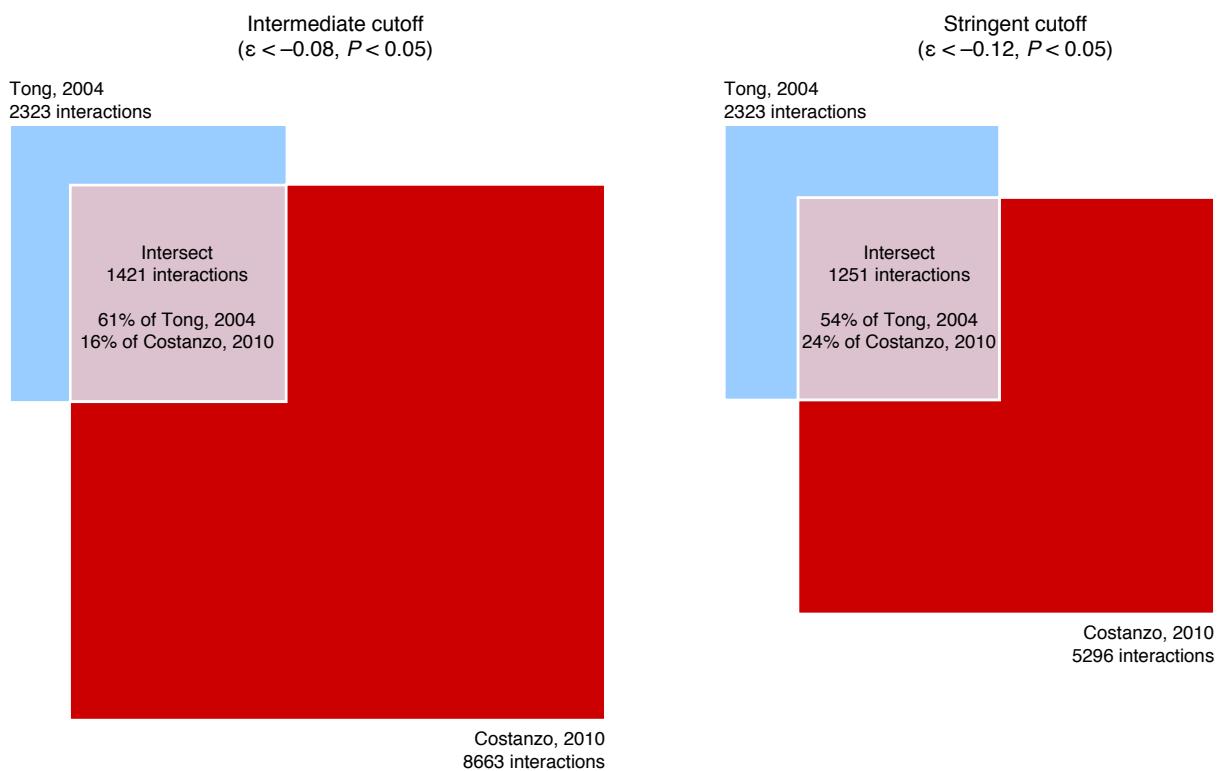
Three arrays were constructed that consisted of the same set of mutant strains located at random plate positions. Colony sizes of the randomized array mutant strains were measured from 10 SGA control screens and compared to colony sizes for the corresponding mutants derived from 10 randomly selected control screens from a large, genome-scale SGA dataset¹. Pearson correlation was used as the similarity metric to compare colony sizes before and after correction. The average correlation and standard deviation across 10 comparisons are shown. Normalization of systematic effects results in a 40% improvement in correlation between colony size measurements for identical mutant strains located at different positions.

Supplementary Figure 3. Reproducibility of reciprocal genetic interactions measured using colony sizes without normalization



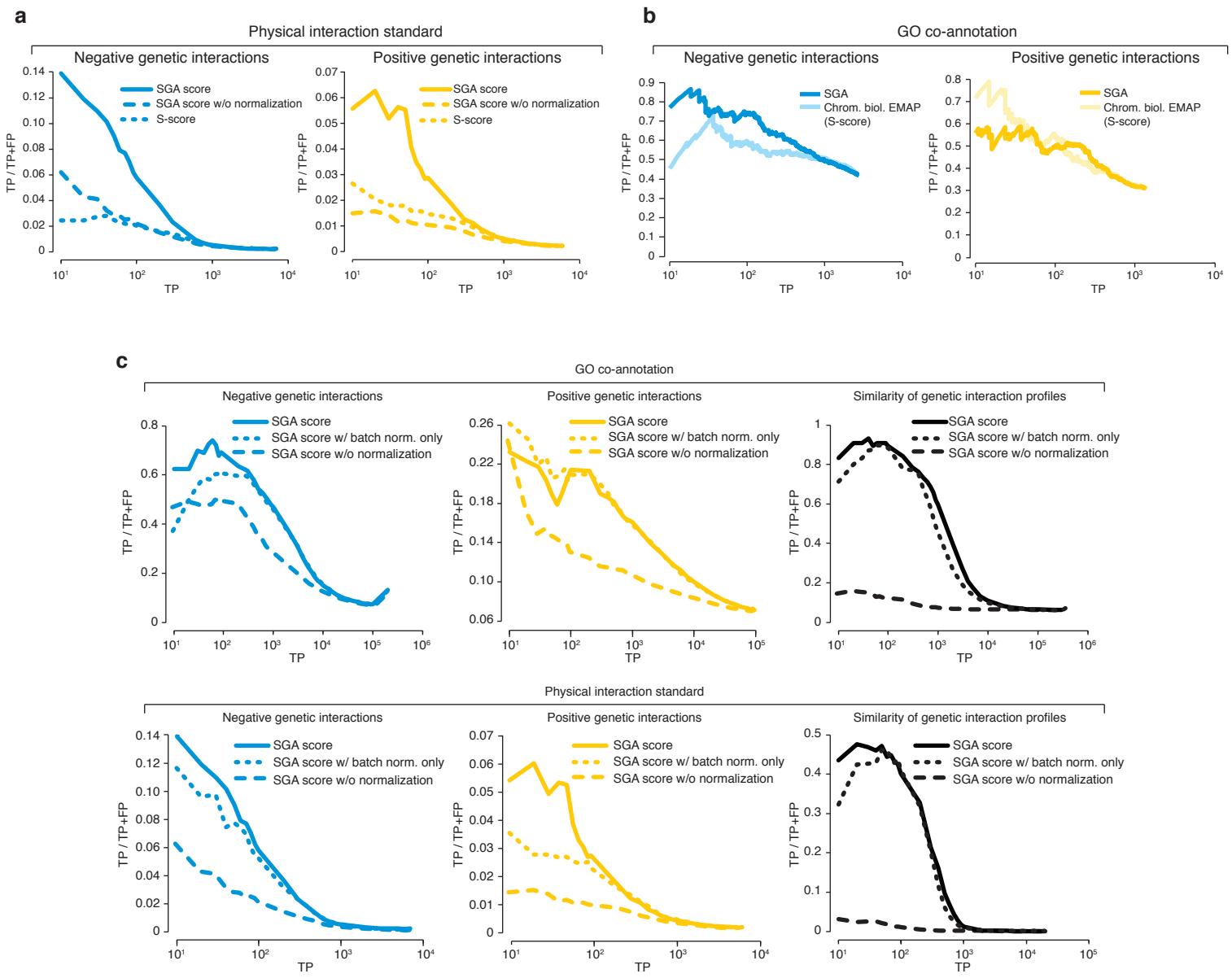
Lenient ($P < 0.05$) and intermediate ($|\varepsilon| > 0.08, P < 0.05$) confidence thresholds were applied.

Supplementary Figure 4. Analysis of genetic interactions reported in a previous SGA study



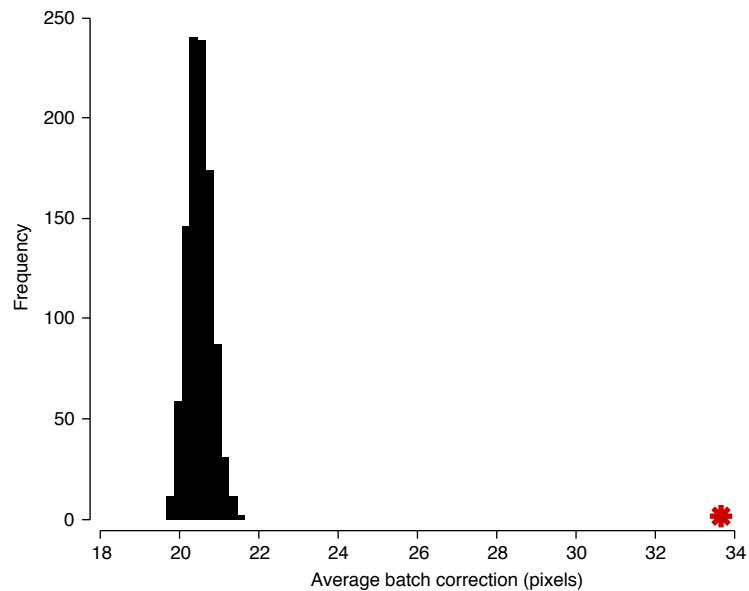
Number of interactions derived from 93 overlapping genetic interaction screens reported in Tong, 2004² and Costanzo, 2010¹. The number of negative genetic interactions as defined by the intermediate and stringent confidence cutoffs in Costanzo, 2010¹ is compared to the number of synthetic sick or lethal interactions from Tong 2004² involving the same query and array genes.

Supplementary Figure 5. Comparative evaluation of negative and positive genetic interactions



(a) Comparative functional evaluation of negative and positive genetic interactions, as determined by the SGA score method applied to normalized colony size data (solid blue and yellow lines), the SGA score method applied to colony size data without normalization (dashed blue and yellow lines) or the S-score method (dotted blue and yellow lines). In all three cases, the same colony size dataset was used. True positive interactions were defined as those involving gene pairs encoding physically interacting proteins (**Online Methods**). Precision and recall were calculated as described previously³. **(b)** Comparative functional evaluation of overlapping genetic interactions from independent studies (SGA data¹ and chromosome biology EMAP data⁴). Only gene pairs present in both the SGA (dark blue and dark yellow lines) and the EMAP (light blue and light yellow lines) datasets were considered. True positive interactions were defined as those involving gene pairs co-annotated to a Gene Ontology gold standard set of terms³. Precision and recall were calculated as described previously³. **(c)** Comparative functional evaluation of negative interactions, positive interactions and genetic interaction profiles to assess the impact of the batch effect. SGA scores and profiles were evaluated using normalized data (solid blue/yellow/black lines), data without normalization (dashed blue/yellow/black lines) or data in which only the batch effect normalization was applied (dotted blue/yellow/black lines). Pearson correlation was used to compute profile similarity for every pair of array mutant strains across 1,712 query mutant strains. True positive interactions were defined as those involving co-annotated gene pairs or protein-protein interactions (**Online Methods**). Precision and recall were calculated as described previously³.

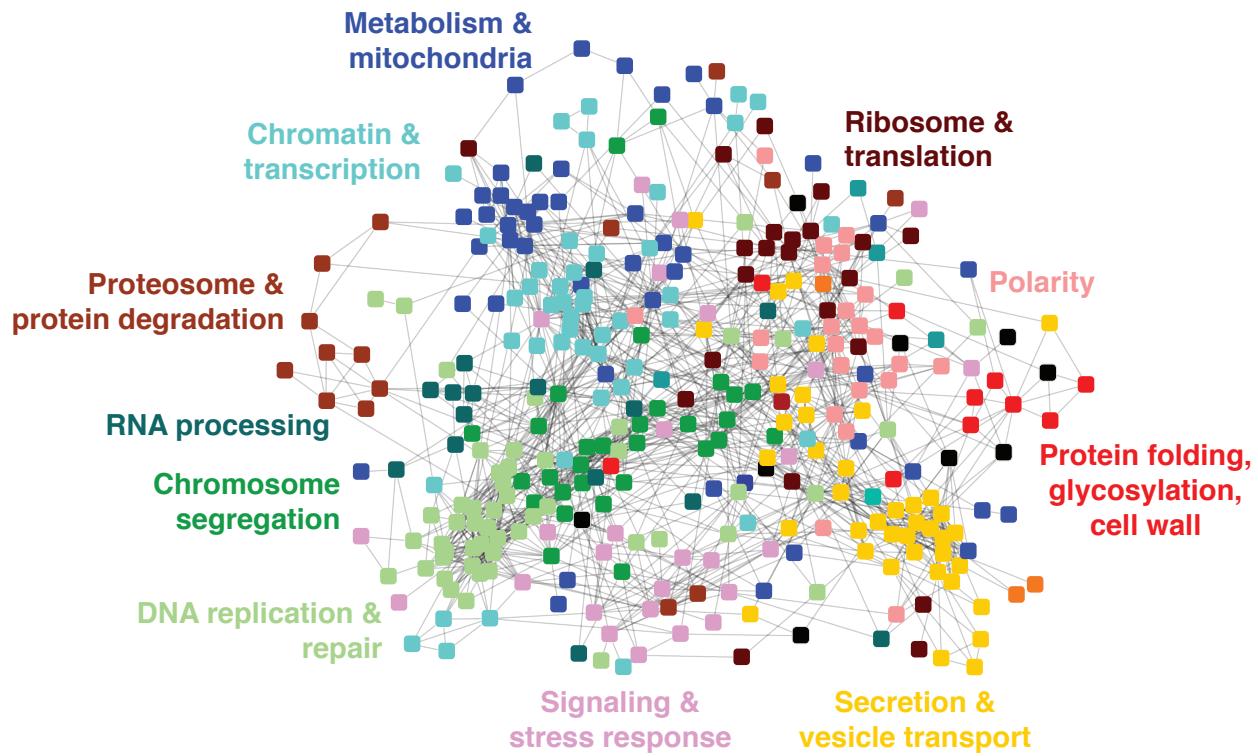
Supplementary Figure 6. Batch effect contribution to the difference between SGA score and S-score functional performance



We computed the average batch correction (in pixels) after normalization for 1,000 gene pairs that show the most extreme negative S-scores but do not appear in the list of the top 1,000 strongest negative SGA score based on the same dataset¹. The average normalization factor for the 1,000 unique S-score interactions is indicated by a red star. For comparison, we randomly selected 1,000 sets of 1,000 gene pairs and plotted the distribution of the average normalization factor (black bars). Negative interactions uniquely identified by the S-score method tend to have more severe batch effects as determined by the SGA score method.

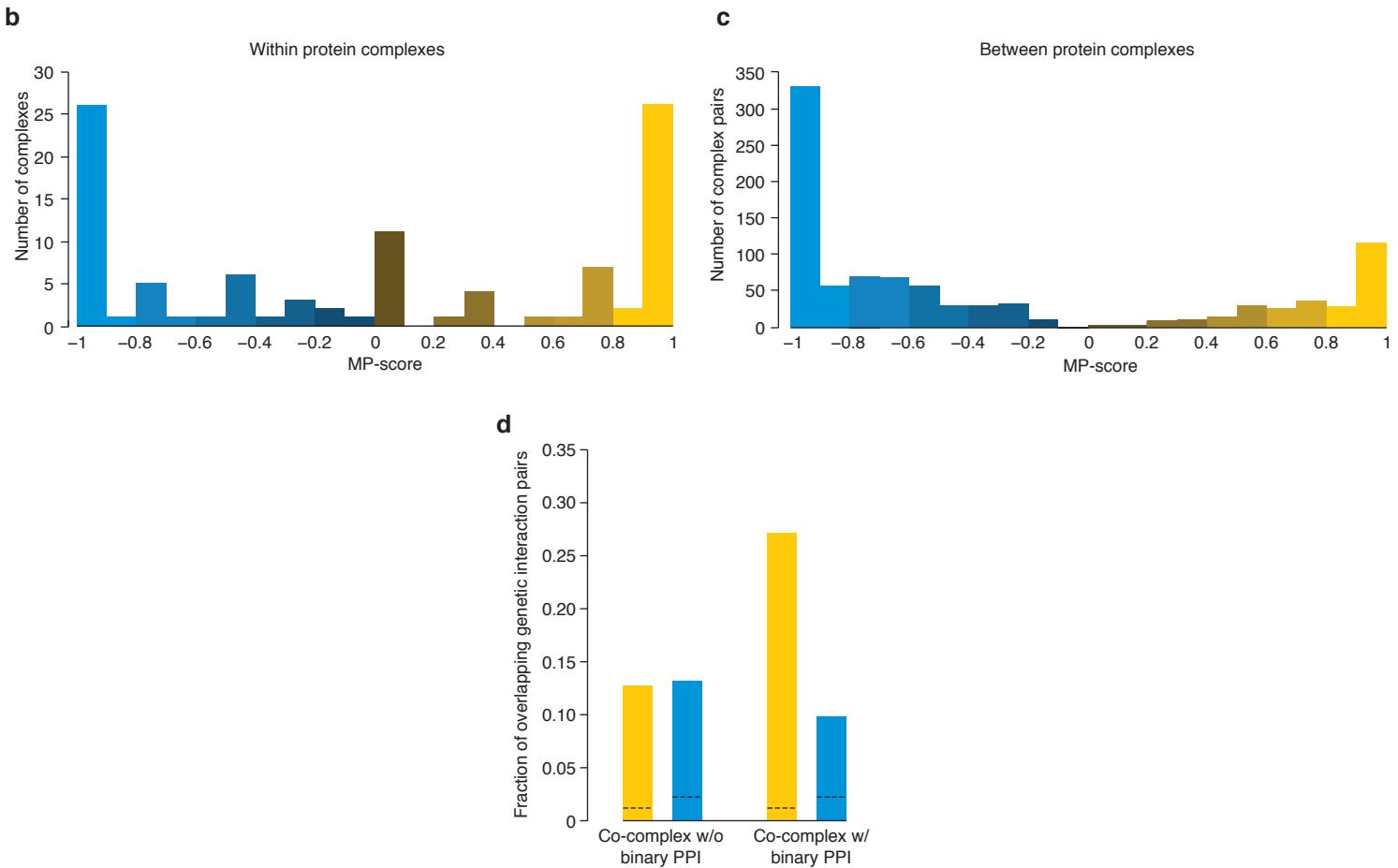
Supplementary Figure 7. Analysis of genetic interactions within and between protein complexes

a



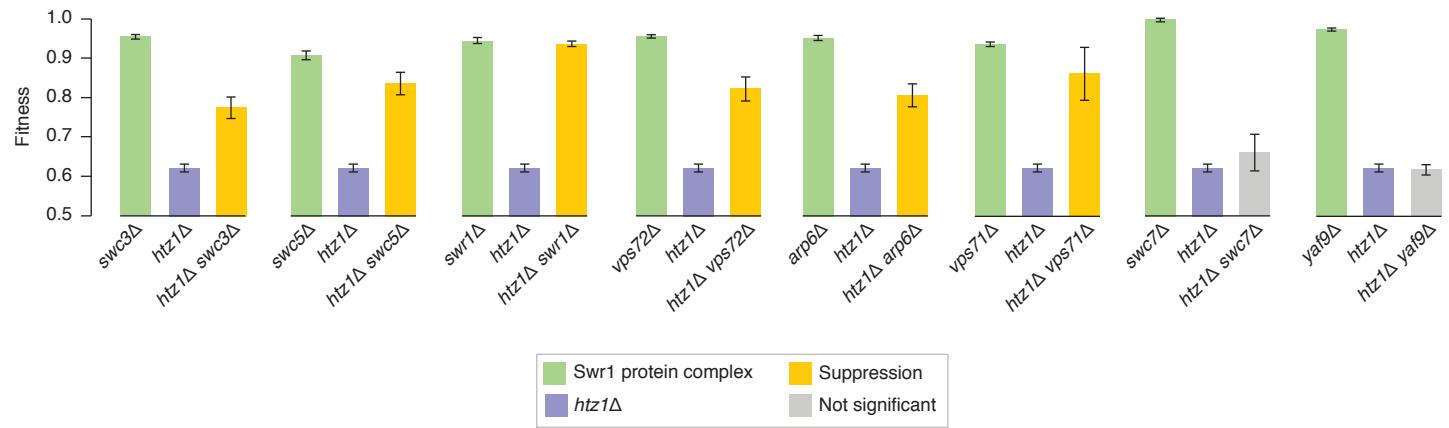
a) A functional map of protein complexes. A module-level network was constructed based on genetic interactions between protein complexes. Genetic profile similarities were measured for all gene pairs by computing Pearson correlation coefficients (PCC) based on complete genetic interaction profiles and averaging across proteins involved in the same protein complex. Network edges were weighted according to the average PCC and an edge weighted, spring embedded layout algorithm was applied to determine node position⁵.

Supplementary Figure 7. Analysis of genetic interactions within and between protein complexes (continued)



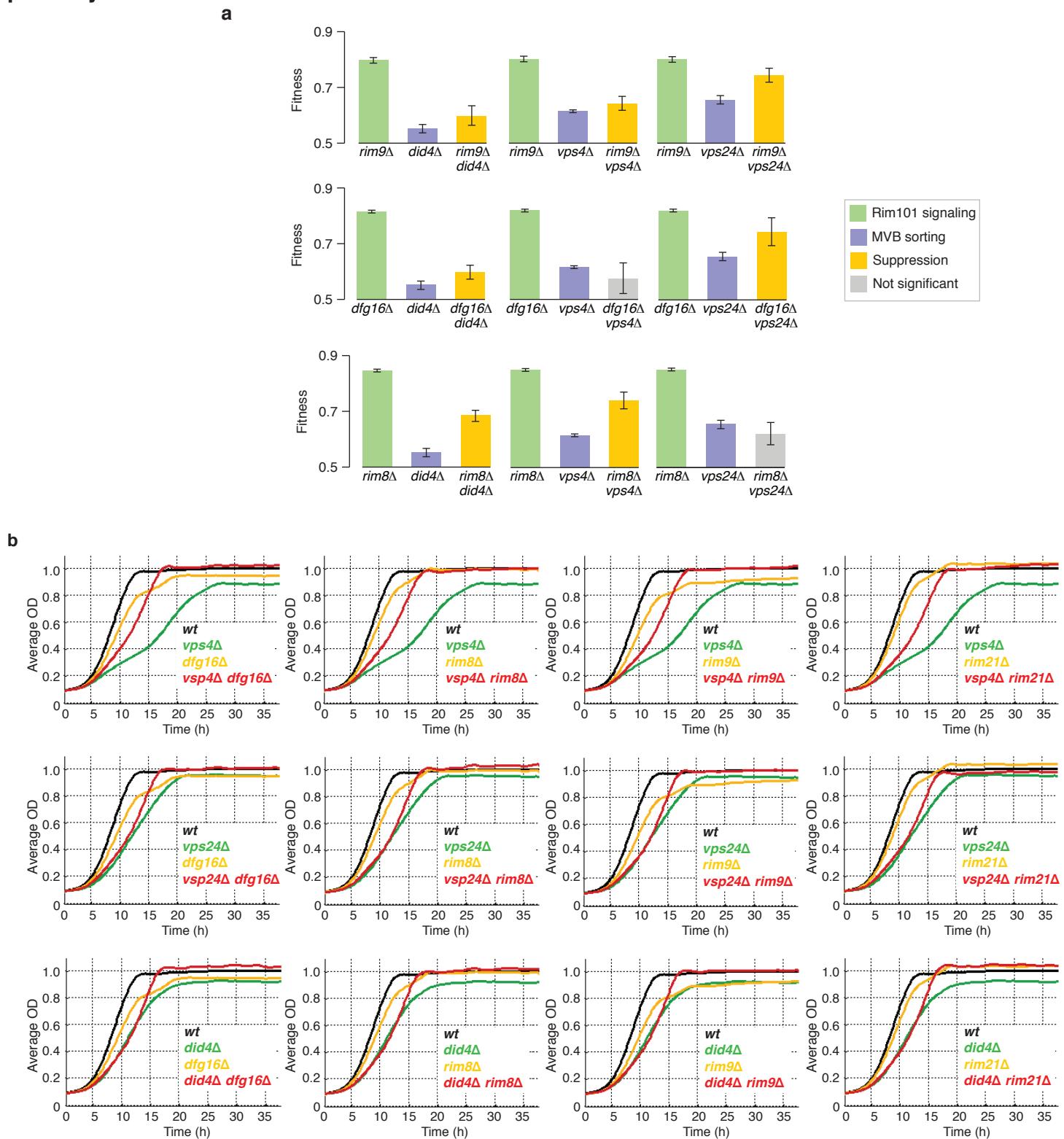
(b) Distribution of complexes with respect to within-complex monochromatic purity scores (MP scores; **Supplementary Note 5**). Only complexes having at least two genetic interactions based on the lenient cutoff ($P < 0.05$) are considered here. **(c)** Distribution of complex-complex pairs with respect to between-complex monochromatic purity scores (MP scores; **Supplementary Note 5**). Edges between complexes are retained when the pair is enriched for genetic interactions (FDR = 5%). **(d)** Genetic interaction frequency among co-complex members with and without a direct protein-protein interaction. Gene pairs within complexes annotated to our protein complex standard were separated into two groups based on whether they also exhibited evidence for a direct protein-protein interaction as detected in a recent two-hybrid-based study⁶. For each group, the fraction of pairs exhibiting either positive (yellow) or negative (blue) genetic interactions was measured after applying an intermediate confidence threshold ($|I| > 0.08$, $P < 0.05$). The background rate of positive and negative interactions among random gene pairs at the same confidence threshold is indicated by the black lines.

Supplementary Figure 8. Genetic suppression between Swr1 protein complex and *htz1Δ*



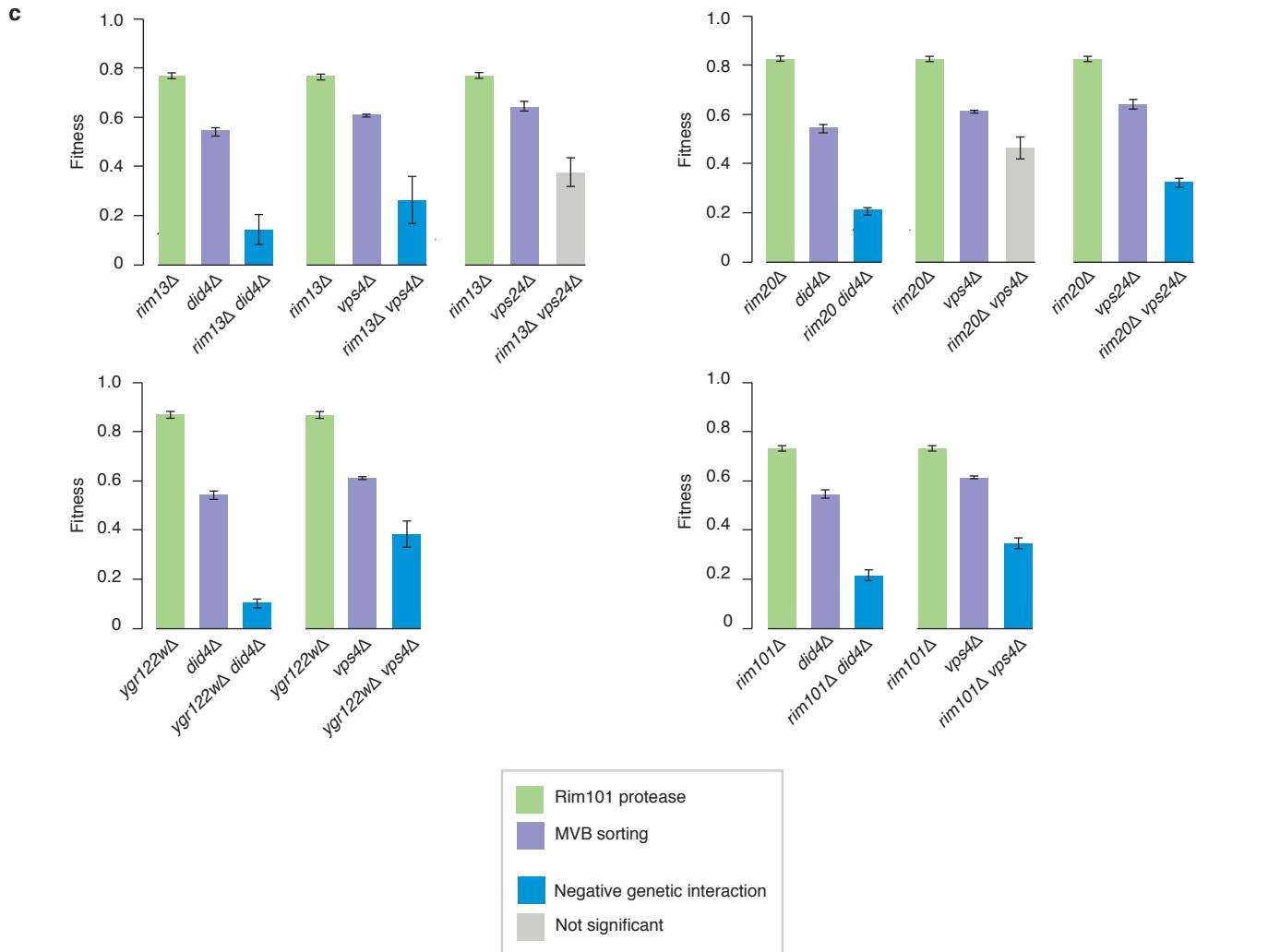
Comparison of colony size-derived single and double mutant fitness measures suggests that mutations in genes encoding members of the Swr1 protein complex (*swc3Δ*, *swc5Δ*, *swr1Δ*, *vps72Δ*, *arp6Δ* and *vps71Δ*) suppress growth defects associated with deletion of *HTZ1*.

Supplementary Figure 9. Genetic suppression between Rim101 signaling and MVB sorting pathways



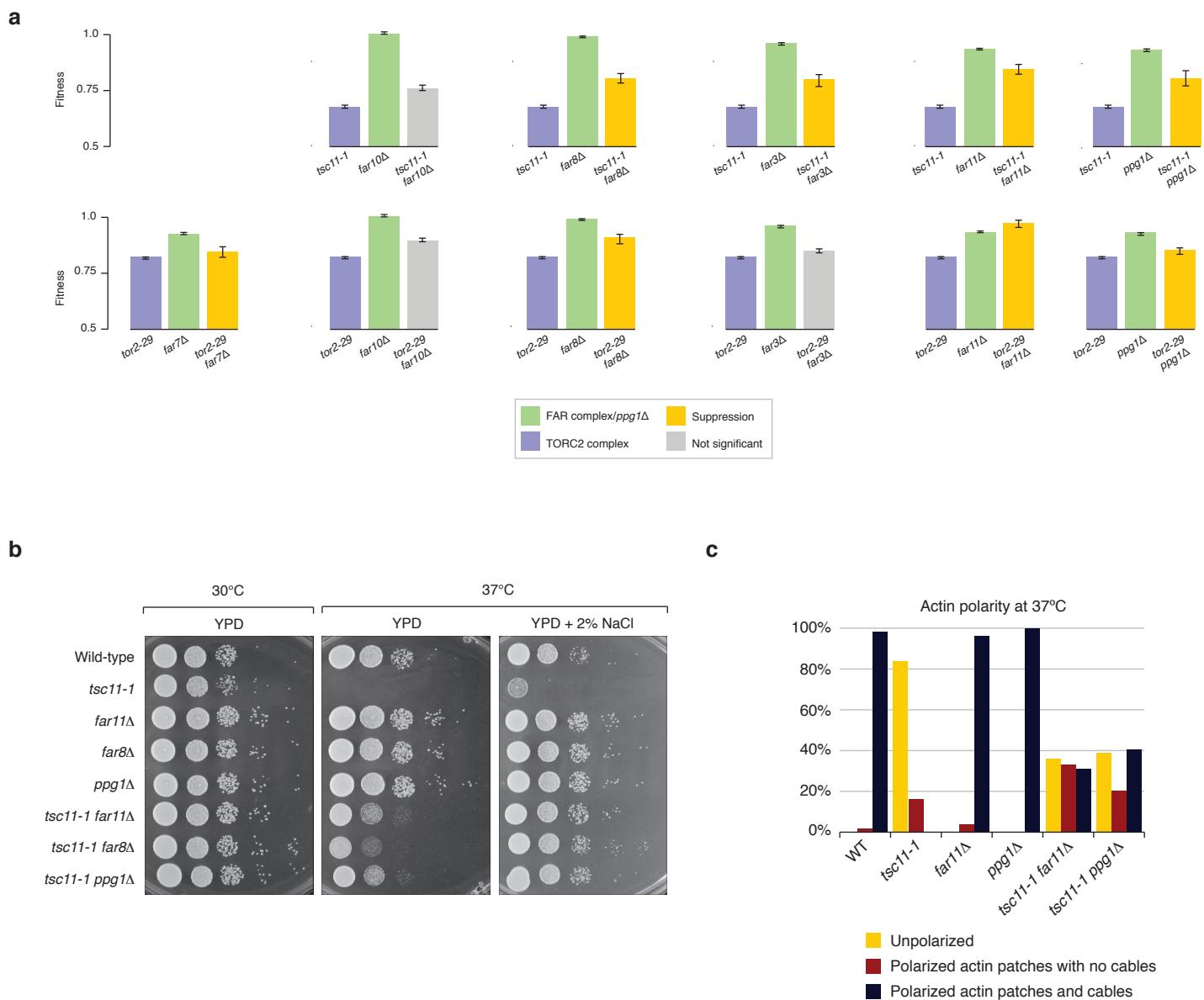
(a) Comparison of colony size-derived single and double mutant fitness measures suggests that *rim8Δ*, *rim9Δ* and *dfg16Δ* suppress growth defects associated with deletion of *DID4*, *VPS4* or *VPS24*. Error bars reflect the s.d. of each single and double mutant fitness. (b) Growth of the wild-type and the indicated single and double mutant strains was monitored in rich media (YPED) as described previously⁷. Each growth curve represents the average of three independent experiments.

Supplementary Figure 9. Genetic suppression between Rim101 signaling and MVB sorting pathways (continued)



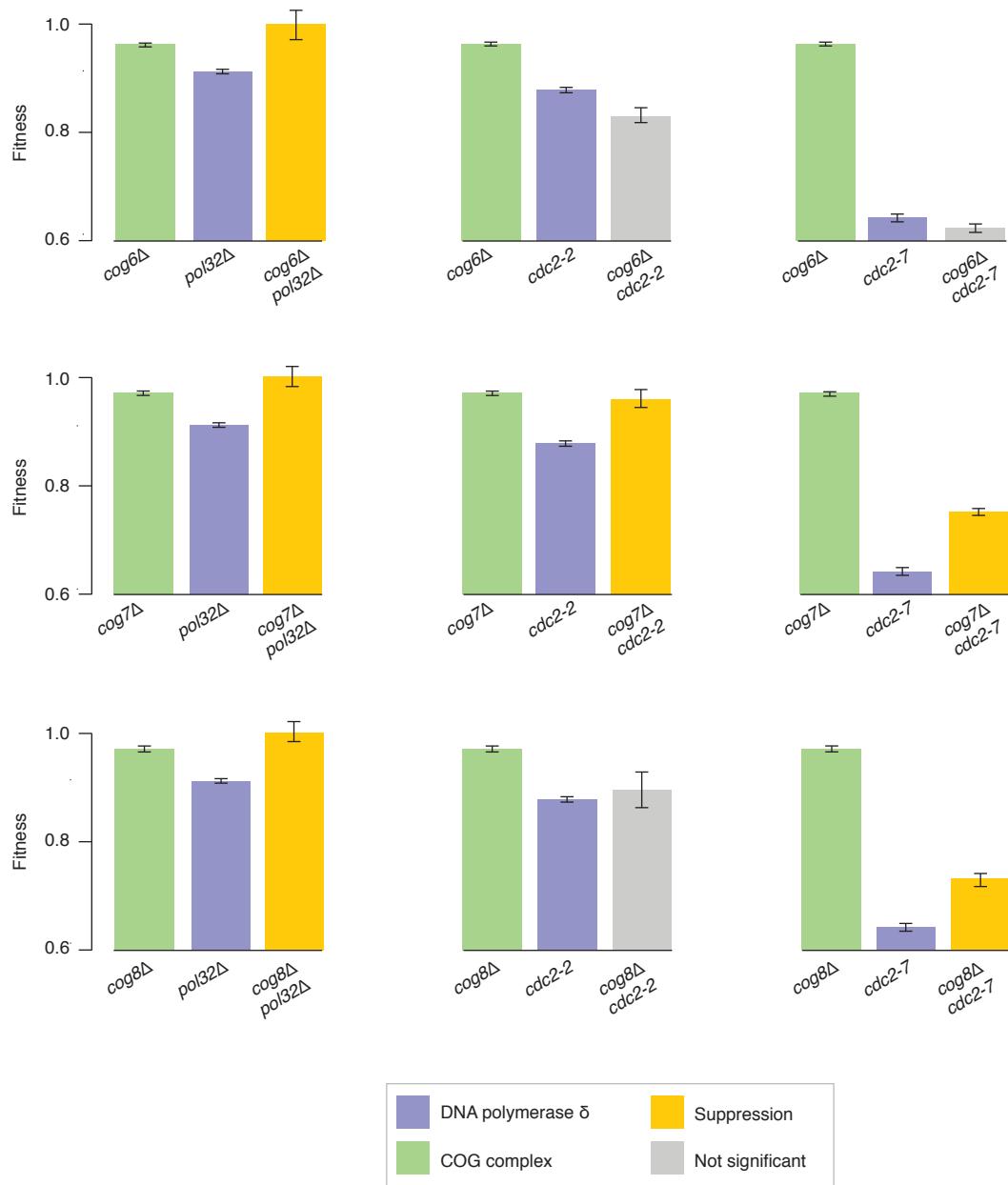
(c) Comparison of colony size-derived single and double mutant fitness measures suggests that deletion of *RIM13*, *RIM20*, *YGR122W* or *RIM101* do not suppress growth defects associated with loss of *DID4*, *VPS4* or *VPS24*.

Supplementary Figure 10. Genetic suppression between FAR and TORC2 protein complexes



(a) Comparison of colony size-derived single and double mutant fitness measures suggests that mutations in genes encoding members of the Far3-11 protein complex suppress growth defects associated with mutant alleles of TORC2 kinase complex gene members, *tor2-29* and *tsc11-1*. Similarly, *tor2-29* and *tsc11-1* mutants are also suppressed by deletion of *PPG1*, a PP2A-related serine/threonine phosphatase. (b) Serial dilution growth assays confirm that *tsc11-1* growth defects are suppressed by deletion of *FAR11*, *FAR8* and *PPG1* under semi-permissive (30°C) and non-permissive (37°C) temperatures (**Online Methods**). Suppression interactions are enhanced in the presence of 2% (0.4 M) NaCl. (c) A *tsc11-1* temperature-sensitive mutant exhibits an abnormal actin morphology that is suppressed by loss of *FAR11* and *PPG1*. The extent of actin polarization was quantified in wild type and the indicated single and double mutants.

Supplementary Figure 11. Genetic suppression between Pol δ and COG protein complex



Comparison of colony size-derived single and double mutant fitness measures suggests that mutations in genes encoding members of the COG protein complex suppress growth defects associated with mutant alleles of DNA polymerase δ genes, *pol32 Δ* , *cdc2-2* and *cdc2-7*.

Supplementary Table 1. List of normalizations

	S-score ⁸	SGA score (this study)
Plate-specific effect normalization	Yes	Yes
Row/column effect normalization	Yes	Yes
Spatial effect normalization	No	Yes
Competition effect normalization	No	Yes
Batch effect normalization	No	Yes

Supplementary Table 2. List of experimental controls

This table lists the most important aspects of the SGA experimental design that play an essential role in the normalization procedures we describe.

Experimental design property	Description
Number of plates per array	The collection of non-essential deletion mutants is spread across 14 agar plates.
Number of colonies and strains per plate	Each array plate consists of 384 mutant strains pinned in quadruplicate for a total of 1,536 colonies/plate, arranged on a grid containing 32 rows and 48 columns. Mutant strain replicates are located next to each other on the same plate.
Single mutant fitness	Single mutant fitness was obtained by crossing an SGA query strain harboring the <i>natMX</i> marker at a neutral locus (<i>URA3</i>) against the non-essential deletion mutant collection. Following SGA, every array <i>kanMX</i> -marked mutant array also contains at <i>natMX</i> marker at the <i>URA3</i> locus. Thus, colony sizes derived from these control screens (after applying standard normalization procedures) reflect array strain single mutant fitness. Fitness measurements were based on 80 control screens. Query strain single mutant fitness was obtained in a similar manner. In this case, the <i>natMX</i> -marked SGA query strains were arrayed and, crossed to a

	single array harboring the <i>kanMX</i> marker at the <i>HO</i> locus via SGA. Colony sizes derived from these control screens (after applying standard normalization procedures) reflect single mutant fitness of the query strains.
Plate border control	Colonies located in the outermost rows and colonies have access to excess nutrients and as a result grow significantly larger than the average colony in the middle of a plate (Supplementary Fig. 1). These border effects are extreme and can be difficult to normalize. To minimize these effects, colonies in the first and last two rows and columns of each array plate (304 colonies) correspond to the same control strain which is isogenic to BY4741 but where the <i>HIS3</i> gene is replaced with the <i>KanMX</i> marker. Histidine auxotrophy is complemented following mating to the SGA query strain which contains the <i>S. pombe his5</i> functional ortholog under the control of a <i>MATa</i> -specific promoter. Border controls are excluded from genetic interaction analysis.
SGA screen reproducibility	A subset of 211 randomly chosen query strains were screened multiple times to assess reproducibility of colony size measurements and genetic interactions (Figs. 2c and 3a).
Reciprocal interactions	Although our current interaction matrix is not complete, we have measured interactions for ~580,000 reciprocal mutant pairs (i.e. queryA-arrayB and queryB-arrayA). We used this set of reciprocal interactions as both a quality control metric as well as a standard for estimating the false positive rate in our dataset.

Supplementary Table 3. Sensitivity and precision of SGA genetic interaction scores

	Negative interactions			Positive interactions		
	Num. interactions	Sensitivity	Precision	Num. interactions	Sensitivity	Precision
Lenient $P < 0.05$	366,085	0.41	0.27	323,935	0.26	0.20
Intermediate $ \varepsilon > 0.08$, $P < 0.05$	108,417	0.35	0.63	59,887	0.18	0.59
Stringent $\varepsilon < -0.12$, $P < 0.05$ $\varepsilon > 0.16$, $P < 0.05$	58,508	0.28	0.89	6,185	0.05	~1

To calculate the number of interactions, reciprocal interactions, where query A was crossed to array B and query B was crossed to array A, were processed as follows. If genetic interaction scores for the double mutants AB and BA show opposite interaction signs (AB is positive and BA is negative, or vice-versa), both pairs were removed. If genetic interaction scores for AB and BA show the same interaction sign (both positive or both negative), the interaction with the lowest p-value was retained. Sensitivity and precision were calculated as described previously¹.

Supplementary Note 1: SGA genetic interaction score

The source code for normalization procedures and the raw colony size data are available as **Supplementary Software 1** and **Supplementary Data 4**. The details of the methods are described below.

Modeling double mutant colony growth

We model colony size as a multiplicative combination of double mutant fitness, time, and experimental factors. Specifically, for a double mutant with deletions of genes i and j , colony area C_{ij} can be expressed as:

$$C_{ij} = \alpha \cdot f_{ij} \cdot t \cdot s_{ij} \cdot e \quad \text{Eq. 1}$$

where f_{ij} is the fitness of the double mutant, t is time, s_{ij} is the combination of all systematic factors, α is a constant scale factor, and e is log-normally distributed error. In addition, double mutant fitness f_{ij} can be expressed as $f_{ij} = f_i f_j + \varepsilon_{ij}$, where f_i and f_j are the single mutant fitness measures, and ε_{ij} is the genetic interaction between genes i and j ⁹. This model is motivated by two empirical observations: (1) replicate plates grown different amounts of time can be normalized by a single multiplicative factor, suggesting that colony area scales linearly with time, and (2) the same array plate crossed into two different query mutants can be normalized by a single multiplicative factor, which suggests colony area also scales linearly with the fitness of each mutation. We also find that most experimental factors tend to scale the colony size, which motivates a multiplicative model.

Our goal is to fit this model and extract the biological factors of interest, which are the single mutant fitnesses f_i and f_j , and the genetic interaction score ε_{ij} . This task is challenging because the systematic experimental factors s_{ij} account for a large proportion of the observed variation in colony size.

Overview of systematic effects

The key obstacle in deriving precise estimates of single and double mutant phenotypes is the presence of several experimental factors that introduce systematic

biases in the observed colony size. We identified several contributors to colony size variance, ranging from spatial artifacts due to how colonies are arrayed, to nutrient competition effects, to experimental batch effects. A summary of all sources of systematic variation and the corresponding normalization schemes are presented below.

Plate effect normalization

A critical step in obtaining accurate double mutant fitness measures is proper normalization across plates. Plate-to-plate variance can be substantial due to the variability in each plate's incubation times as well as to the fact that overall plate colony size is related to the corresponding query mutation. From **Eq. 1** above, we assume that observed colony area is a function of the single mutations fitness defects as well as time:

$$C_{ij} = \alpha \cdot i_j \cdot f_j \cdot t \cdot s_{ij} \cdot e \quad \text{Eq. 2}$$

as in most cases $\varepsilon_{ij} \approx 0$ because genetic interactions are rare. Due to the experimental design (see **Supplementary Table 2** for details), a single plate consists of a single query, j , crossed into an array of 384 array single mutants pinned in quadruplicate. Each plate is grown until the average colony size reaches a minimum level, as judged by the technician, and plates are only imaged once. Thus, the time each plate is grown is linked directly to the fitness of the query and their product is relatively constant, i.e. $f_j t \approx c$ for some constant, c . Through empirical studies, we found it difficult to separately estimate f_j and t , and thus, we simply normalize out the product of the two (and any other plate-related effects). We follow a normalization scheme described previously⁸, and estimate a plate normalization factor by computing the plate middle mean (PMM , mean of the middle 60% of colonies on the plate). All colonies on a plate k are then normalized as follows:

$$\tilde{C}_{ij} = C_{ij} \frac{PMM_{global}}{PMM_k} \quad \text{Eq. 3}$$

where PMM_{global} is derived from the PMM across all plates. Note that $PMM_k \approx f_j t_k$ where t_k is the growth time for plate k . This process completely removes any dependence

of the normalized colony size on the query mutation in the absence of a genetic interaction ($\varepsilon_{ij} \approx 0$). The new normalized colony size can now be expressed as:

$$\tilde{C}_{ij} = \alpha \cdot f_i \cdot \tilde{s}_{ij} \cdot e \quad \text{Eq. 4}$$

When an interaction is present, the normalized colony size depends on the query fitness as follows:

$$\tilde{C}_{ij} = \alpha \left(f_i + \frac{\varepsilon_{ij}}{f_j} \right) \tilde{s}_{ij} e \quad \text{Eq. 5}$$

As noted in⁸, these normalized colony sizes represent a convenient space for measuring interactions. To detect an interaction, the colony of interest can simply be compared to the background of all queries crossed into that particular array strain, which is in the same position across all plates. After plate normalization, the double mutant in each position should reflect only the array mutation's single mutant fitness and the associated systematic effects (i.e. $\tilde{C}_{ij} \approx \alpha f_i \tilde{s}_{ij}$) unless a genetic interaction is present.

Interactions for a particular double mutant can then be estimated by simply comparing the colony size for the corresponding array to the colony size in exactly the same plate position across the entire collection of screens (queries). As described in⁸, this essentially provides an empirical “expected” fitness of the double mutants corresponding to that array strain because the effect of the query mutation has been normalized out across all screens. Following this logic, we define a colony residual R_{ij} as

$$R_{ij} = \tilde{C}_{ij} - \text{median}_j \tilde{C}_{ij} \approx \alpha \frac{\varepsilon_{ij}}{f_j} \tilde{s}_{ij} e \quad \text{Eq. 6}$$

which reflects the interactions ε_{ij} we hope to measure (in addition to lingering systematic effects, \tilde{s}_{ij}). This is simply the difference between the normalized colony size for a particular double mutant and the median normalized colony size across all query strains that have been screened against the corresponding array strain. Note that we could also define a relative version of the colony residual by measuring the ratio (or log-ratio) of the

normalized colony of interest to the median colony size in that position. We will revisit this choice later in the context of normalizing specific systematic effects.

There are several advantages to normalization on the colony residual space as opposed to the original colony space. First, the variance in the normalized colony space reflects both the biological variation due to single mutant fitness of the array strain (f_i), positional systematic biases, as well as the variation of interest, genetic interactions. Variation on the colony residual measures, however, is orthogonal to the variation due to the array strain single mutant fitness, and furthermore, several positional biases are common among plates, and thus differences between the same positions across multiple plates are less susceptible to positional biases. Observed variation on the colony residual is more directly related to the quantity of interest, the degree of the interaction between two genes, and thus normalization on this space is more effective than normalization on colony size itself. The normalization procedures described below are mainly based on colony residuals, except where noted.

Normalization of row/column effects

Similar to other array-based genomic technologies (**Supplementary Note 2**), spatial effects are a major contributor to systematic variation in colony sizes. This is immediately apparent when one looks at imaged double mutant colonies (**Supplementary Fig. 1**). For example, colonies in rows or columns near the borders of the plate are visibly larger, which is mainly due to the reduced colony density near the plate edges and the resulting availability of extra nutrients. The estimation of row/column effects is done on a plate-specific basis using a linear LOWESS smoothing¹⁰ on the normalized colony sizes. While the shape of the row/column biases tends to be reproducible, it can be more or less severe from plate to plate, and thus normalization is applied on a plate-by-plate basis. Estimating these effects is challenging given a limited number of unique strains per row/column. Because these effects are due to the geometric arrangement of colonies on the plate and the corresponding availability of nutrients, we expect that neighboring rows/columns should exhibit similar effects, and consequently, trends across or down the plate should be relatively smooth. We can take advantage of this property to derive accurate estimates of how colony position affects colony size and

remove this systematic trend from the data. Specifically, we apply linear LOWESS smoothing on normalized colonies (**Eq. 7**) to estimate the colony size-row and colony size-column trends, using window size spanning six rows or columns. This smoothing typically follows the trend one might expect based on visual inspection of plates: colonies in the outer rows and columns tend to grow larger (**Supplementary Fig. 1b**). Interestingly, we also identify other subtle yet statistically significant trends such as an overall W-shape and a slight increase in colony size as one moves across and down the plate. The trends are highly consistent across plates, suggesting they are a real systematic artifact, likely due to the geometric properties of the plate or the media. Row and column correction factors are applied as follows:

$$\tilde{C}_{ij}' = \tilde{C}_{ij} \left(\frac{\bar{r}}{r_j} \right) \left(\frac{\bar{c}}{c_k} \right) \quad \text{Eq. 7}$$

where r_j and c_k are estimates derived from LOWESS smoothing for row j and column k and \bar{r} and \bar{c} denote average row and average column factors.

Normalization of local spatial gradients

Many array plates exhibit a spatial gradient pattern in colony size that is often more local than the row/column effect described above and can be highly variable from plate to plate (**Supplementary Fig. 1c**). We suspect this is due to variation in the thickness of media across the plate, which becomes significant when one considers residual difference in colonies at the same position. Consequently, this effect can result in apparent interactions on the same order as the most extreme real genetic interactions (**Supplementary Fig. 1c**). Spatial gradients are estimated based on colony residuals using a series of 2D smoothing filters. First, a median spatial filter is applied on the 7×7

grid of colonies surrounding a position of interest using $\log\left(\frac{\tilde{C}_{ij}}{\text{median}_j \tilde{C}_{ij}}\right)$. The median

filter estimates are then further smoothed with a simple average filter with on a 10×10 grid to derive final estimates of the surface gradient. The spatial normalization is applied as follows:

$$\tilde{C}_{ij}^{'} = \tilde{C}_{ij} \frac{1}{\exp(s_{ij})} \quad \text{Eq. 8}$$

where s_{ij} is the spatial effect estimated by the spatial filters.

Nutrient competition effects and normalization

Another substantial systematic effect that hinders our ability to measure accurate genetic interactions is local competition for nutrients between neighboring colonies. This effect is largely due to the high density of colonies on the plate (1,536 total per plate), and is most pronounced in cases where a healthy colony is positioned next to a sick colony or an empty spot. The severity of this effect is illustrated in **Supplementary Fig. 1d**. The result of this factor is that the largest double mutant colonies typically are associated with small neighbors, suggesting that the reason they are large is not that they are more fit, but that they have access to more nutrients in the local neighborhood. If not corrected, this effect can result in thousands of spurious positive genetic interactions (**Supplementary Fig. 1d**).

To normalize the nutrient competition effect, we take a two-step approach. First, all colony residuals are binned into ten deciles based on their neighbor colony sizes. Normalization is then applied to remove the effect of competition both within and between these groups as described below.

To normalize the competition effect within each decile, we plot double mutant colony residual versus the minimum neighbor size and apply linear LOWESS smoothing with a window size of 500. The within-group effect is normalized by scaling colonies by the results of the LOWESS smoothing at the corresponding neighbor size.

The competition effect is normalized across decile groups by using quantile normalization, which is a technique that has been applied extensively in the area of microarray normalization¹¹. Essentially, quantile normalization takes the data one wants to normalize and a reference distribution, and forces the cumulative distribution of the sample data to match the cumulative distribution of the reference data. We define a reference distribution based on double mutants with relatively healthy neighbors (60–80

percentile), and then quantile normalize each decile described above such that the cumulative distribution function matches this reference.

Screen batch effect and normalization

Another experimental factor that critically affects our ability to derive precise single and double mutant phenotypes from mutant colony growth is the batch effect. We refer to “screens” as a single query mutant that has been crossed into all plates of arrayed single mutants (e.g. a whole-genome screen for one query). We find that screens completed together (by the same technician using the same robot on the same day) tend to share a common “batch signature”. Grouping by batch can be caused by a number of different factors. For example, the media for these screens was all from the same batch and plates poured in exactly the same way, the same technician prepared the query lawn at the beginning of the SGA protocol, the plates were grown for a similar amount of time, or the screens used a common source array for the robotic pinning steps. The combination of these factors results in a signature that consists of a set of unusually small or large colonies that would otherwise look like extreme negative or positive interactions. This effect is particularly problematic if one hopes to use profiles of genetic interaction to predict function of screened genes. Genes screened in the same experimental batch tend to share a common (non-biological) signature, which is often stronger than the real biological interaction signature (**Supplementary Fig. 1e**). Removing the batch effect is statistically challenging because batches typically consist of 5–10 screens while the batch signature can involve 1,000–10,000 colonies. Estimating such a high-dimensional signature from a limited number of examples faces the same curse of dimensionality that has posed statistical challenges for related effects in microarrays (**Supplementary Note 2**)^{12–14}.

Linear discriminant analysis (LDA) is an approach used for finding linear combinations of features that are highly discriminative across a set of classes. LDA is related to the commonly used principal component analysis (PCA), or more generally, singular value decomposition (SVD), but is a supervised method that is able to leverage pre-defined class distinctions, which are experimental batches in our context.

Specifically, if we assume a matrix, X , of double mutant colony data, where rows of X consist of screens grouped into different batches, the objective of multi-class LDA is to find projections of the data that maximize the following criterion:

$$J(v) = \frac{\det(v^T S_b v)}{\det(v^T S_w v)} \quad \text{Eq. 9}$$

where

$$S_b = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)$$

$$S_w = \sum_{i=1}^C \sum_{x_k \in C_i} (x_k - \mu)(x_k - \mu)$$

for C different batches where μ_i is the batch i centroid and μ is the global centroid. Intuitively, this criterion measures the ratio of between-batch differences to within-batch differences and the optimal vector v represents linear combinations of the features that maximize this batch signal-to-noise ratio. Optimal projections v can be readily computed by finding the principal eigenvectors of the matrix $S_w^{-1}S_b$. Principal eigenvectors identified through LDA are effectively focused towards cross-batch variation, which should avoid removing potentially important (biological) signatures from the interactions.

To normalize the batch effect, we construct a matrix, V_n consisting of the top n principal eigenvectors and normalize the original data matrix as follows:

$$X' = X - X * V_n^T V_n \quad \text{Eq. 10}$$

This computes the projection of the original colony size data onto the space of “batch signals” and removes these signatures from the data. The number of components removed is chosen as the principal eigenvectors satisfying $\lambda_i/\lambda_{\max} > 0.1$ for eigenvalues λ_i .

For all screens, batches are defined by clustering time-stamps on images of each plate since these time-stamps will be similar for screens processed together. Such groups usually consist of 5–10 query screens where each screen consists of 14 different plates.

Batch normalization is critical for deriving precise interactions and, in particular, for using interaction profiles to predict function. As discussed above, before batch normalization, the average correlation coefficient for two screens in the same batch was as high as the typical correlation due to biological co-function (**Supplementary Fig. 1e**). After normalization, the same-batch correlation drops dramatically, and consequently, our ability to predict function from similar interaction profiles improves (**Fig. 4**).

Additional normalization and colony data processing procedures

Removing colonies influenced by physical linkage of genetic markers

Genetically linked genes (genes located close to one another on the same chromosome) tend to show negative genetic interaction scores because recombination frequency between these genes is reduced and, thus, the number of meiotic progeny containing both the query and the array mutations is also reduced. To filter out these spurious negative interactions, we map the linkage region corresponding to each query mutation by identifying the set of adjacent negative interactions overlapping the chromosomal location of the query gene. Genetic interaction scores were smoothed along the chromosome using a window size of 7. The linkage region was defined as the largest region such that smoothed interaction scores around the genomic region of interest remained negative. Using this approach, the size of the linkage region that is removed from consideration varies based on the specific query gene. The mean linkage region is 410 kb with a standard deviation of 150 kb.

Large colony filter

Due to problems with some strains forming diploids during the SGA protocol selection steps, we filtered any double mutants showing abnormally large colonies from our analysis. Specifically, any double mutants for which all four colonies were greater than 1.5 times the median colony size were filtered.

Jackknife variance filter

Due to various technical issues (e.g. contamination or pinning errors during robotic manipulation) in some cases, one colony in each group of replicate double mutants on the same plate would grow considerably larger or smaller than its replicates. To filter these colonies, we applied a jackknife procedure to assess the impact of removing each colony on the estimated variance. For each double mutant, all replicate colonies were collected, and the standard deviation was estimated leaving each replicate colony out one at a time. Individual colonies, i , were filtered when $\sigma - \sigma_i > (2/N) \cdot \sigma$ where σ is the mean jackknife standard deviation estimate, σ_i is the jackknife estimate derived from holding out replicate i , and N is the total number of replicates.

Order of filtering and normalization steps

Raw colony areas were derived from segmented images and the processing steps described above were applied serially in the following order: linkage filter, plate normalization, large colony filter, spatial gradient normalization, row/column normalization, competition correction, jackknife variance filter, and batch normalization.

Fitness and interaction estimation procedure

The fitness and interaction estimation procedure consists of a series of normalization steps (described above) followed by two independent steps for estimating the biological factors of interest: the single and double mutant fitness.

Recall that in the absence of genetic interactions, the normalized colony sizes directly reflect the array single mutant fitness. Thus, fitness estimates can be readily derived for any single mutant that appears on the array by simply computing the median across all occurrences of that array:

$$\text{median}_j \tilde{C}_{ij}' = \text{median}_j \alpha \left(f_i + \frac{\varepsilon_{ij}}{f_j} \right) e \approx \alpha f_i \quad \text{Eq. 11}$$

Unlike the estimation of interactions, which is most sensitive to variance on the colony residual, precise estimation of single mutant fitness depends critically on proper normalization at the colony size level. For example, positional biases consistent across plates do not present a problem for detecting interactions, but could lead to highly inaccurate estimates of the array single mutant fitness, f_i . To ensure precise estimation of single mutant fitness, we constructed several array configurations, with all non-essential single mutant deletion strains occurring in multiple different spatial contexts. We applied all normalization procedures to these screens and computed the median over all control screens for each array strain to derive single mutant fitness estimates for all non-essential genes. The parameter α was estimated by setting the mode of the fitness distribution to 1, reflecting our assumption that most deletions should have near wild-type fitness.

Variance in single mutant fitness estimates was estimated from bootstrap sampling of the median, and final fitness estimates were derived by pooling across each array spatial configuration. A detailed analysis of the quality of these estimates in comparison with single mutant phenotypes from other assays is discussed in the Results section (**Fig. 2**).

Quantitative estimates of genetic interactions and double mutant fitness phenotypes are derived from normalized colony residuals following the normalization steps. Specifically, for each replicate colony k of a double mutant sharing mutations (or deletions) in genes i and j we compute the colony residual

$$R_{ijk} = \tilde{C}_{ijk}' - \text{median}_j \tilde{C}_{ijk}' \quad \text{Eq. 12}$$

where \tilde{C}_{ijk}' represents the colony size after all normalization steps. An interaction estimate is derived then by averaging across the N_{ij} replicate colonies (typically four per screen with up to two screens) as follows:

$$I_{ij} = \frac{1}{N_{ij}} \sum_k R_{ijk} \quad \text{Eq. 13}$$

where f_j and α are derived as discussed in the previous section.

Another critical aspect of measuring interactions is deriving a corresponding accurate estimate of variance. For each measured interaction, we associate two different variance measures: one that reflects the local variability in replicate colonies (based on 4–8 colonies for the corresponding double mutant), and another pooled estimate that reflects the expected variability of the double mutant in question. The first variance measure is derived from the unbiased estimate of variance on the interaction mean:

$$\sigma_{I_{ij}} = \sqrt{\frac{\sum_k (R_{ijk} - I_{ij})^2}{N_{ij} - 1}} \quad \text{Eq. 14}$$

As reported in previous work on quantitative analysis of SGA screens⁸, variance estimates derived from local double mutant colony reproducibility can occasionally dramatically underestimate the true variance. We also noted this issue in our study and attribute the observation to the fact that replicate colonies are situated next to each other on the plate, and thus, experience much of the same systematic variation (i.e. they are not independent). To derive a more accurate variance estimate, we pooled error estimates across all occurrences of a given array strain and query strain for the double mutant of interest and combined them to derive a baseline “expected” variance. From **Eq. 2** above, recall our assumption that, in the absence of interaction ($\varepsilon_{ij} \approx 0$), colony size is proportional to the product of single mutant fitness estimates:

$$C_{ij} \propto \alpha f_i f_j \quad \text{Eq. 15}$$

Thus, we assume that in the null case of no interaction, colony size is log-normally distributed with variance contributions from both the query and the array mutations:

$$C_{ij} \sim \alpha f_i f_j e^X \quad \text{Eq. 16}$$

$$\text{for } X \sim N(0, \sigma_i^2 + \sigma_j^2)$$

where σ_i^2 and σ_j^2 are array strain and query strain-specific variance contributions. We estimated the array strain-specific variance contribution by measuring variance across several wild-type control screens taken through all processing and normalization steps.

These screens were constructed to lack any biological interactions, and thus, any observed variation in a fixed array-plate position across plates, could be attributed to array strain variance. We computed query-strain specific variance by pooling within-array variance estimates across all array strains for the corresponding query screen. For each measured interaction, we report both the measured standard deviation on the interaction estimate itself, $\sigma_{I_{ij}}$, and the multiplicative geometric standard deviation,

$$\sigma_{ij_{\text{expected}}} = e^{\sqrt{\sigma_i^2 + \sigma_j^2}}, \text{ based on the expected combination of array and query variance.}$$

These variance estimates can be used, for example, to estimate a worse-case confidence interval on an interaction estimate.

Post-interaction filters

In our experimental setup, array mutant strains are distributed across 14 plates in chromosomal order, such that on average each individual plate contains mutant strains from one to three chromosomes. Due to infrequent meiotic recombination between linked genes, plates harboring array mutant strains that are genetically linked to a given query mutation contain an abnormally large number of inviable or slow growing colonies. These slow-growing colonies generate extreme competition effects that are not completely removed by our normalization procedures. Thus, to avoid introducing spurious positives genetic interactions, we filter any positive genetic interaction located adjacent to the query strain linkage group.

Supplementary Note 2: Statistical comparison of SGA genetic interaction screens to gene expression microarrays

Commonality with microarray setting:

- (1) The raw data obtained from genetic interaction screens depends on segmentation of colonies arrayed in an ordered grid, which is similar to the image segmentation/intensity measurement process used in microarray analysis.
- (2) Both microarrays and genetic interaction screens produce high dimensional, quantitative read-outs for each screen. This presents many of the same issues for normalization of the resulting data. For example, an experimental batch in SGA might have 4–5 screens while each screen has ~4,000 quantitative measurements. This is similar to the dimensionality of expression arrays, where batches consistent of ~5 or less arrays, each with tens of thousands of quantitative expression measures across the genome.
- (3) Both microarrays and SGA screens exhibit array-to-array, plate-to-plate and position-to-position differences.
- (4) Both technologies are highly susceptible to batch effects. Arrays that are run as a group or interaction screens experience similar experimental treatment, which ultimately leads to common systematic biases. This, combined with a high-dimensional, quantitative read-out, can make cross-array or cross-screen comparisons challenging.

Differences from the microarray setting:

- (1) The size of the grid is significantly smaller in the genetic interaction setting (1,536 colonies per plate as compared to tens of thousands spots on a typical microarray slide).
- (2) The causes and the nature of the spatial variation can be quite different between the two technologies. For example, in the context of microarrays, air pockets under the cover slip, scratches, or other technical issues such as uneven washing during or after hybridization can introduce spatial effects. In addition to similar technical issues in the context of SGA screens (e.g. query strain lawn spread unevenly, colony contamination across positions because of technician/robot error), nutrient competition is

a major factor in colony array-based screens. For example, the size of the colony at one position determines the availability of nutrients nearby, and so spatial effects are sometimes directly related to, or introduced by, the quantity of interest (colony size) in neighboring positions.

(3) A typical microarray study has a much smaller number of arrays than a large-scale application of SGA has screens. For example, the data analyzed here includes ~1,700 whole-genome screens, whereas the largest array studies typically include tens or possibly hundreds of arrays. This difference has important implications on possibilities for normalizing various experimental artifacts, for example, the batch effect. A typical genetic interaction dataset has hundreds of different batches compared to only a handful (<10) in a microarray setting, which enables different normalization approaches.

(4) The magnitude of the biological effect being measured in genetic interaction screens relative to systematic variation is much smaller than in the context of gene expression analysis. Genetic interactions are rare—even subtle effects are expected to occur in less than 5% of the mutants screened. Thus, only a small fraction of the observed variation in mutant colony size is actually related to the biological effect of interest. This is quite different in the microarray setting, where in a particular condition one might expect significant expression changes in 20–30% of the genome. This also has important implications for the types of normalization procedures that can be applied.

Supplementary Note 3: Reciprocal interaction analysis

As noted in the “Evaluating genetic interaction measurements” section, one of the approaches used to measure the reproducibility of genetic interactions derived with the SGA score method was to compute the correlation of reciprocal interactions. For example, the interaction derived from query mutant A crossed into array mutant B should be similar to the interaction derived from query mutant B crossed into array mutant A. As noted in the main text, this correlation is $r = 0.67$, which is notably lower than correlation derived from other measures of reproducibility, including two replicate screens ($r = 0.87$) or the correlation with liquid growth-derived interactions ($r = 0.89$). There are a number of experimental reasons why reproducibility is lower by this metric as compared to the

other approaches. First, reciprocal interactions A-B and B-A are measured under entirely different experimental conditions (e.g. different screen, different position on plate, different experimental batch). Second, and more importantly, each SGA screen is adjusted to maximize the dynamic range for the given query mutant. For query mutants with larger fitness defects, the entire set of array plates is grown longer, which generally allows for greater resolution in detecting genetic interactions. Thus, screens including the reciprocal interactions A-B or B-A may have different sensitivities depending on the relative fitnesses of the A or B single mutants, which can lead to reduced correlation. This can result in certain interactions (either positive or negative) being detected in one screen (query A crossed into array B), which are not as easily detected in the reciprocal screen (query B crossed into A). The fact that the correlation of reciprocal interactions is lower than replicate screens is reflective of this difference in detection sensitivity.

Supplementary Note 4: Comparison to genetic interactions curated in BioGRID

We compared positive and negative interactions produced by our approach to those reported in previous genetic interaction studies, as curated in BioGRID¹⁵. For negative genetic interactions, we considered BioGRID interactions annotated as phenotypic enhancement, synthetic growth defect and synthetic lethality. For positive genetic interactions, we considered BioGRID interactions annotated as phenotypic suppression and synthetic rescue. The overlap with these interactions, combined with the agreement of reciprocal gene pairs, enabled an independent estimation of precision and sensitivity at various confidence thresholds (**Supplementary Table 3**). For example, at an intermediate interaction cutoff ($|\varepsilon| > 0.08$, $P < 0.05$), we estimate a precision of 0.63 and a sensitivity of 0.35 for negative genetic interactions (**Supplementary Table 3**).

Supplementary Note 5: Evaluating overlap of genetic and protein-protein interaction networks

Construction of a protein complex standard

A literature-curated protein complex standard was compiled by combining the two most recent protein complex standards available for yeast. This standard consists of 430 complexes derived from SGD Macromolecular Complex GO standard (www.yeastgenome.org), CYC2008 protein complex catalog¹⁶ and 26 manually curated complexes/pathways. Redundant protein complex annotations were minimized by eliminating all but one complex with identical components and by excluding smaller complexes if all their members also belong to a larger complex. Partially overlapping complexes were treated as separate complexes. A list of protein complexes and their members is provided in **Supplementary Data 2**.

Analysis of genetic interactions within and between protein complexes

For all analyses of genetic interactions within and between protein complexes, we used interactions at the lenient cutoff ($P < 0.05$) to maximize coverage of physically interacting pairs.

Enrichment analysis of genetic interactions within protein complexes

Each complex for which at least two pairs were screened for genetic interactions was assessed for enrichment of either positive or negative interactions among its members. We ignored all dubious ORFs and considered the union of genetic interactions for each gene with multiple alleles screened. Significance was evaluated using the hypergeometric distribution as follows:

$$P = 1 - \sum_{n=0}^{X-1} \frac{\binom{K}{n} \binom{M-K}{N-n}}{\binom{M}{N}}$$
 Eq. 17

where

M = total number of screened gene pairs in this study

K = total number of gene pairs with positive/negative genetic interactions

N = total number of screened gene pairs within protein complex

X = total number of interacting (positive/negative) gene pairs within protein complex

Monochromatic analysis of genetic interactions within protein complexes

To evaluate the monochromaticity of interactions within protein complexes, we measured the purity of interactions for any complex with enrichment for interactions by the criteria described above ($P < 0.05$). Specifically, we defined a monochromatic purity score (MP-score) as follows:

$$MP(C_i) = \frac{\frac{1}{N_i} \sum_{j,k \in C_i} e_{jk} - \alpha}{1 - \text{sign}\left(\frac{1}{N_i} \sum_{j,k \in C_i} e_{jk}\right) * \alpha} \quad \text{Eq. 18}$$

where

$$\alpha = \frac{1}{N_{tot}} \sum_{\forall j,k} e_{jk}$$

N_i = number of screened pairs within complex i

N_{tot} = number of total screened pairs

C_i = all possible proteins (genes) in complex i

$$e_{jk} = \begin{cases} +1 & \text{if genes } j \text{ and } k \text{ have a positive GI} \\ -1 & \text{if genes } j \text{ and } k \text{ have a negative GI} \end{cases}$$

This score is designed such that a complex with pure positive genetic interactions will have an MP-score equal to +1, while a complex having all negative pairs will have an MP-score equal to -1. A complex reflecting the background ratio of positive to negative interactions will present a MP-score equal to 0. Monochromatic complexes

described in **Fig. 5** are complexes satisfying $|MP(C_i)| > 0.5$. A distribution of all complexes monochromatic purity scores is shown in **Supplementary Fig. 7b**.

Construction of a complex-complex network of genetic interactions

To enable a study of interactions at the module level, we defined a complex-complex network of genetic interactions (**Fig. 5b**). Complex-complex pairs were assessed for enrichment of genetic interactions as well as monochromaticity. Enrichment of interactions within complex-complex pairs was assessed as follows:

$$P = 1 - \sum_{n=0}^{X-1} \frac{\binom{K}{n} \binom{M-K}{N-n}}{\binom{M}{N}}$$
Eq. 19

where

M = [total number of screened partners outside from complex i] U [total number of screened partners outside from complex j]

K = [total number of genetic interaction partners outside from complex i] U [total number of genetic interaction partners outside from complex j]

N = [total number of screened pairs between complex i and j]

X = [total number of genetic interaction pairs between complex i and j].

Between-complex interaction analysis was restricted to positive and negative interactions identified at the lenient cutoff ($P < 0.05$). Monochromaticity was assessed similarly to the approach used for within-complex monochromaticity (described above) but based on all gene-pairs spanning across each pair of complexes. The complete distribution of between-complex MP scores is shown in **Supplementary Fig. 7c**. For the complex-complex network degree analysis in **Fig. 5b**, a between-complex edge in the modular network was conservatively assumed to exist where the pair had significant enrichment (false discovery rate 5% based on hyper-geometric distribution and Westfall and Young step-down procedure for multiple hypothesis correction) and the between-complex MP score satisfied $|MP(C_i-C_j)| > 0.75$. The complex-complex network degree of

purely negative ($\text{MP-score}_{\text{within}} = -1$) and purely positive ($\text{MP-score}_{\text{within}} = 1$) complexes was measured and compared (**Fig. 5b, inset**), and the Wilcoxon rank-sum test was used to assess the significance of the difference in degree between complexes connected by pure negative and pure positive genetic interactions.

Supplementary Note 6: Protein complex suppression network

To construct a complex-complex suppression network, we first identified all complex-complex pairs with significant enrichment for positive interactions and a minimum of five shared positive genetic interactions (**Supplementary Note 5** and **Supplementary Data 3**). We then categorized the between-complex positive interactions into suppression and masking subclasses. Assuming that f_a , f_b and f_{ab} are the fitness measures for single mutants a , b and the double mutant ab , respectively, and that $f_a = \max(f_a, f_b)$, we used the following rules: (1) if $f_{ab} > f_b + \sigma_b$ (where σ_b is the standard deviation of f_b), mutant a was determined to suppress the phenotype of mutant b ; (2) if rule (1) was not true, but if $f_{ab} < f_a - \sigma_a$, then mutant b was determined to mask the phenotype of mutant a . Prior to this analysis, all positive interactions were filtered in the following way: a) $\varepsilon > 0.08$; b) $P < 0.05$; c) $f_a, f_b < 1$; d) $|f_a - f_b| \geq \sigma_a + \sigma_b$. We identified all complex-complex pairs with greater than 80% directional and suppression consistency among the corresponding gene pairs and visualized them as a network (**Fig. 6a**).

Supplementary References

1. Costanzo, M., Baryshnikova, A., Bellay, J. & al, e. The Genetic Lanscape of a Cell. *Science* **303**, 808-813 (2010).
2. Tong, A.H. et al. Global mapping of the yeast genetic interaction network. *Science* **303**, 808-813 (2004).
3. Myers, C.L., Barrett, D.R., Hibbs, M.A., Huttenhower, C. & Troyanskaya, O.G. Finding function: evaluation methods for functional genomic data. *BMC Genomics* **7**, 187 (2006).
4. Collins, S.R. et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**, 806-810 (2007).
5. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
6. Yu, H. et al. High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104-110 (2008).
7. St Onge, R.P. et al. Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat Genet* **39**, 199-206 (2007).
8. Collins, S.R., Schuldiner, M., Krogan, N.J. & Weissman, J.S. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol* **7**, R63 (2006).
9. Elena, S.F. & Lenski, R.E. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395-398 (1997).
10. Cleveland, W.S. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74** (1979).
11. Bolstad, B.M., Irizarry, R.A., Astrand, M. & Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193 (2003).
12. Johnson, W.E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118-127 (2007).
13. Benito, M. et al. Adjustment of systematic microarray data biases. *Bioinformatics* **20**, 105-114 (2004).
14. Leek, J.T. & Storey, J.D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* **3**, 1724-1735 (2007).
15. Breitkreutz, B.J. et al. The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res* **36**, D637-640 (2008).
16. Pu, S., Wong, J., Turner, B., Cho, E. & Wodak, S.J. Up-to-date catalogues of yeast protein complexes. *Nucleic Acids Res* **37**, 825-831 (2009).