

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

Anastasia Baryshnikova^{1,2,10}, Michael Costanzo^{1,10}, Yungil Kim^{3,4}, Huiming Ding¹, Judice Koh¹, Kiana Toufighi¹, Ji-Young Youn^{1,2}, Jiongwen Ou⁵, Bryan-Joseph San Luis¹, Sunayan Bandyopadhyay³, Matthew Hibbs⁶, David Hess⁷, Anne-Claude Gingras⁸, Gary D Bader^{1,2}, Olga G Troyanskaya⁹, Grant W Brown⁵, Brenda Andrews^{1,2}, Charles Boone^{1,2} & Chad L Myers³

Global quantitative analysis of genetic interactions is a powerful approach for deciphering the roles of genes and mapping functional relationships among pathways. Using colony size as a proxy for fitness, we developed a method for measuring fitness-based genetic interactions from high-density arrays of yeast double mutants generated by synthetic genetic array (SGA) analysis. We identified several experimental sources of systematic variation and developed normalization strategies to obtain accurate single- and double-mutant fitness measurements, which rival the accuracy of other high-resolution studies. We applied the SGA score to examine the relationship between physical and genetic interaction networks, and we found that positive genetic interactions connect across functionally distinct protein complexes revealing a network of genetic suppression among loss-of-function alleles.

Synthetic genetic array (SGA) analysis is an automated form of yeast genetics that combines arrays of mutant strains with robotic manipulations for high-throughput double-mutant construction¹. Genetic interactions are identified as unexpected phenotypes arising from the combination of two or more genetic variants and include two broad categories: 'positive' and 'negative'². Negative genetic interactions refer to a more severe fitness defect than expected, with an extreme case being synthetic lethality; positive genetic interactions refer to double mutants with a less severe fitness defect than expected. Quantitative assessment of fitness-based genetic interactions requires measurements of single-mutant fitness, an estimate of the expected double-mutant fitness, which is typically modeled as a multiplicative combination of the single-mutant phenotypes, and measurement of the observed double-mutant fitness, such that the difference between the observed and expected double-mutant fitness can be calculated².

Genetic interactions previously have been quantified from SGA experiments of functionally biased subsets of a yeast deletion mutant collection using an interaction score (S score)³. This approach relies on colony-size measurements obtained from digital images to determine a composite score derived from both the magnitude and reproducibility of genetic interactions³. The S score is an estimate of the confidence with which genetic interactions can be assigned but does not reflect single- or double-mutant fitness⁴, which are critical for detailed interpretation of interactions and the resultant network^{5,6}. Furthermore, quantification of colony size measurements using the S score is not as precise as that achieved by higher-accuracy, albeit lower-throughput, methods⁴. Finally, mapping genome-wide, fitness-based genetic interactions requires normalization of several parameters associated with colony-based interaction assays (**Supplementary Table 1**), as these can produce unreliable interactions when applied to whole-genome screens.

We developed the SGA score, which relates mutant colony size to fitness and enables identification of quantitative genetic interactions from high-throughput, genome-scale SGA screens. To do so, we identified several systematic biases associated with genome-scale SGA analysis and developed normalization methods to remove these biases. This led to single- and double-mutant fitness measurements of comparable precision to those obtained using other high-resolution methodologies. We applied the SGA score to a genome-scale collection of SGA screens to quantify genetic interactions among ~5.4 million gene pairs⁷. Here we describe the SGA score and accompanying normalization strategies, and report the results of a global comparative analysis between the genetic network deduced from this analysis and the yeast

¹Banting and Best Department of Medical Research, Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Canada.

²Department of Molecular Genetics, Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Canada. ³Department of Computer Science and Engineering, University of Minnesota–Twin Cities, Minneapolis, Minnesota, USA. ⁴Department of Electrical and Computer Engineering, University of Minnesota–Twin Cities, Minneapolis, Minnesota, USA. ⁵Department of Biochemistry, Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Canada. ⁶The Jackson Laboratory, Bar Harbor, Maine, USA. ⁷Department of Biology, Santa Clara University, Santa Clara, California, USA. ⁸Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada. ⁹Department of Computer Science, Lewis-Sigler Institute for Integrative Genomics, Carl Icahn Laboratory, Princeton University, Princeton, New Jersey, USA. ¹⁰These authors contributed equally to this work. Correspondence should be addressed to B.A. (brenda.andrews@utoronto.ca), C.B. (charlie.boone@utoronto.ca) or C.L.M. (cmymers@cs.umn.edu).

physical interaction network. We found that positive genetic interactions connected functionally distinct protein complexes more frequently than described previously.

RESULTS

Quantitative, fitness-based model for genetic interactions

The requirement for accurate phenotypic measurements has imposed constraints on the scale and functional scope of quantitative genetic interaction studies^{4,5,8}. Genome-scale surveys require experimental designs to optimize throughput and thus often sacrifice accuracy. We developed a strategy for deriving quantitative genetic interactions from arrays of double-mutant yeast colonies applicable to genome-scale SGA screens, such that all possible *Saccharomyces cerevisiae* digenic interactions could be examined in an accurate and unbiased manner.

SGA methodology enables rapid and systematic construction of yeast double mutants by mating a strain harboring a ‘query’ mutation of interest to input arrays of strains carrying different ‘array’ mutations, which are composed either of nonessential deletion mutants or conditional alleles of essential genes^{1,9} (Fig. 1a). After several robot-facilitated selection steps, the final output array consisting of haploid double-mutant colonies are imaged at a single time point^{1,9} (Fig. 1a). We developed a model for relating the area of a double-mutant colony image to the fitness of the constituent single mutants by assuming that, in the absence of genetic interactions, double-mutant fitness is a multiplicative combination of single-mutant fitness and experimental factors (Supplementary Note 1). Then we measured genetic interactions as deviations from the expected double-mutant fitness.

Similar to other high-throughput technologies (Supplementary Note 2), genetic-interaction screens using ordered mutant arrays are susceptible to experimental factors that introduce systematic variation in colony size (Fig. 1b and Supplementary Note 1). Contributing factors include subtle differences in growth conditions, such as duration of incubation, from one array plate to the next³ (Fig. 1b and Supplementary Fig. 1a) as well as factors that influence local nutrient availability and affect growth of different subsets of colonies on the same plate. These include plate

location, gradients in growth medium volume caused by uneven preparation surfaces and neighboring mutant-strain fitness (Fig. 1b and Supplementary Fig. 1b–d). Each factor had a noticeable impact on colony size (Fig. 1b,c and Supplementary Fig. 1f). For example, mutant strains located in the outermost rows and columns were on average 40% larger than centrally located mutants (Supplementary Fig. 1b). Strains next to less-fit mutants, including those showing negative genetic interactions, were also larger, suggesting that local competition for nutrients should be considered carefully, especially when measuring positive genetic interactions (Supplementary Fig. 1d).

The most problematic source of systematic variability was ‘batch effect’ (batch is defined as a set of screens conducted in series using the same robotic instrument; Fig. 1b and Supplementary Fig. 1e,f). Because screens conducted in a batch are influenced by common experimental factors, double mutants in a batch often exhibit similar trends in colony-size variation irrespective of the identity of the query strain (Supplementary Fig. 1e). Colony-size variation owing to batch effects was often similar in magnitude to the genetic interactions we aimed to measure; without accounting for the batch effect, the similarity of the interaction profile of functionally unrelated genes was often as strong as that between known complexed protein pairs (Supplementary Fig. 1e). Given the importance of correlation-based approaches for genetic-interaction analyses⁹, strategies for normalizing batch signatures are critical for harnessing the full potential of large-scale genetic-interaction data.

We developed normalization procedures that estimate and remove systematic biases in colony size arising from experimental factors (Supplementary Note 1). We applied statistical techniques including spatial smoothing, quantile normalization and linear discriminant analysis to substantially reduce colony size variability and improve correlation between independent colony size measurements obtained for the same mutant strain located in different array positions (Fig. 1c, Supplementary Figs. 1f and 2 and Supplementary Note 1). After removal of experimental artifacts, we fit a model estimating fitness and genetic interactions for each double mutant to the normalized colony sizes (Fig. 1c

Figure 1 | The SGA score for measuring quantitative genetic interactions. (a) An SGA experiment crossing a strain carrying a query mutation to an input array of single mutants, each of which carries a wild-type copy of the query gene and a unique array strain mutation. A final output array of double mutants is generated after several SGA selection steps, photographed and processed using software that measures colony areas in terms of pixels. Relative colony size, determined by measuring deviation of individual colonies from the median size for the same colony across 1,712 different experiments⁷, is shown. (b) Schematic depiction of the five factors that contribute to experimental variance of colony size. (c) Relative colony size after normalization. Single-mutant fitness (W_A , W_B) and double-mutant fitness (W_{AB}) derived from normalized colony size measurements were used to identify and measure genetic interactions (SGA score; ϵ).

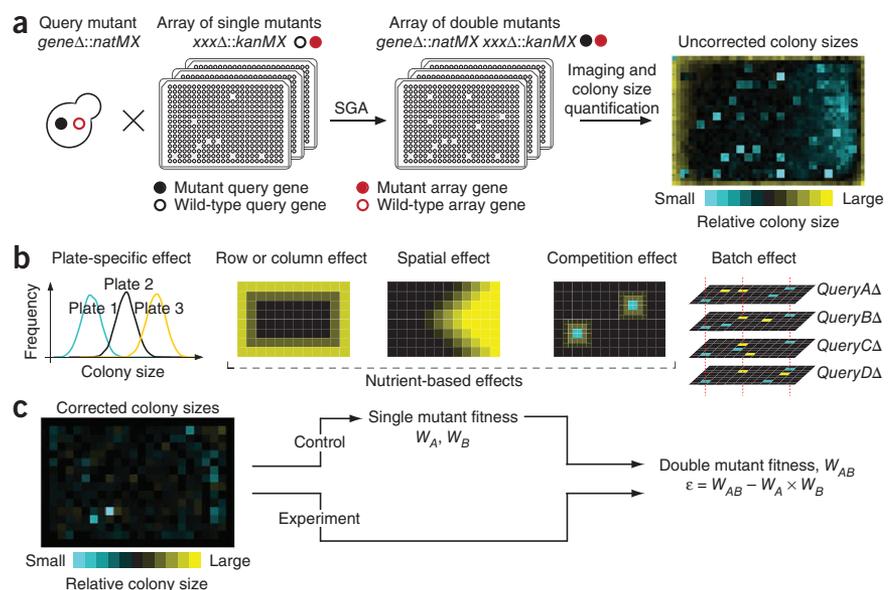
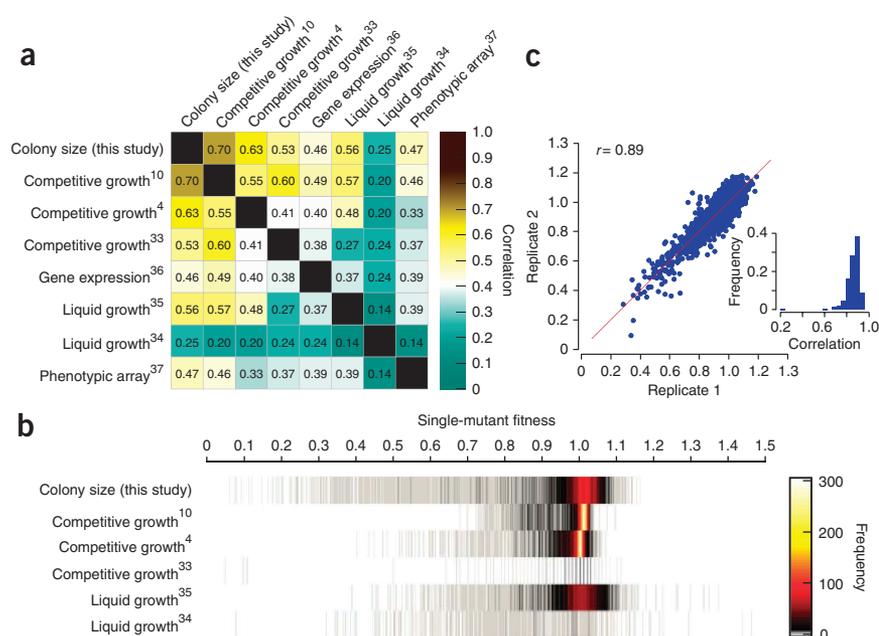


Figure 2 | Evaluation of single-mutant fitness measures. **(a)** Comparison of single-mutant fitness or relative growth measurements for nonessential gene deletions derived from eight independent approaches: colony size measurements (this study), competitive growth analyzed by barcode hybridization¹⁰, flow cytometry⁴ or gene expression profiles³³, gene-expression microarrays³⁶, liquid growth profiling^{34,35} and a spot assay on solid growth medium (phenotypic array)³⁷. **(b)** Distribution of single-mutant fitness measures reported by the studies described in a reporting fitness or relative growth rate. **(c)** Correlation of double-mutant fitness measures obtained from two independent replicates of a representative genome-wide SGA screen. Red line, $y=x$. Inset, distribution of correlations between double-mutant fitness measures obtained from 211 genome-wide SGA screens conducted in duplicate.



and **Supplementary Note 1**). The resulting measure, termed SGA score, captures single- and double-mutant fitness measurements and provides a quantitative, genome-wide assessment of genetic interactions.

A high-resolution catalog of yeast fitness

We measured colony sizes for 4,635 viable deletion mutants and 1,388 temperature-sensitive or hypomorphic alleles of essential genes after SGA analysis using a neutral (control) query mutation (**Fig. 1**, **Supplementary Table 2** and **Supplementary Note 1**). Using image-analysis software, we obtained colony areas from images (Online Methods) and processed them using the SGA score to systematically analyze yeast single-mutant fitness based on colony size (**Supplementary Data 1**). Comparison to previously published fitness or relative growth rate studies revealed that, together with barcode-based fitness measurements of single-deletion strains¹⁰, our colony size-based fitness measures had the highest average cross-study correlation (**Fig. 2a**), suggesting that large-scale colony size measurements capture fitness as well or better than available methodologies. The range of single-mutant

growth defects detectable by our method substantially exceeded that of most other studies. For example, even though growth competition assays assessed a similar number of strains^{4,10}, they had not reported fitness for mutants with relative growth <50% that of a wild-type strain (**Fig. 2b**).

Our colony size-based measurements of double-mutant fitness were highly reproducible as determined by comparative analysis of 211 genome-wide replicate screens (**Fig. 2c**). Therefore, the accuracy afforded by colony size-based single- and double-mutant measurements should provide a basis for quantitative, fitness-based assessment of genetic interactions.

Evaluating genetic-interaction measurements

We used multiple approaches to evaluate the quantitative nature of genetic interactions identified using the SGA score. First, genetic interactions measured from independent experiments performed in duplicate were highly reproducible (**Fig. 3a**). Second, we assessed the reproducibility of interactions identified

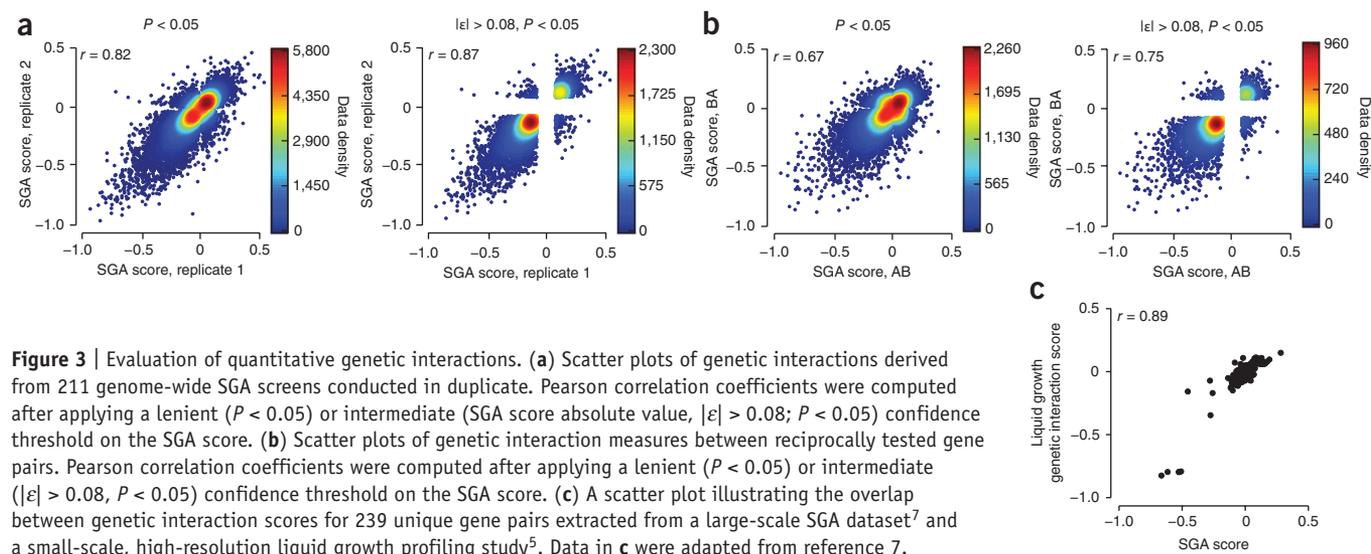


Figure 3 | Evaluation of quantitative genetic interactions. **(a)** Scatter plots of genetic interactions derived from 211 genome-wide SGA screens conducted in duplicate. Pearson correlation coefficients were computed after applying a lenient ($P < 0.05$) or intermediate (SGA score absolute value, $|r| > 0.08$; $P < 0.05$) confidence threshold on the SGA score. **(b)** Scatter plots of genetic interaction measures between reciprocally tested gene pairs. Pearson correlation coefficients were computed after applying a lenient ($P < 0.05$) or intermediate ($|r| > 0.08$, $P < 0.05$) confidence threshold on the SGA score. **(c)** A scatter plot illustrating the overlap between genetic interaction scores for 239 unique gene pairs extracted from a large-scale SGA dataset⁷ and a small-scale, high-resolution liquid growth profiling study⁵. Data in **c** were adapted from reference 7.

among reciprocal gene pairs (query mutant A and array mutant B versus query mutant B and array mutant A). Query mutants, which are mutants crossed into the deletion mutant array, and the corresponding array mutants are subject to different experimental conditions that affect their colony growth differently in the SGA assay. In particular, query mutants with larger fitness defects are allowed to grow longer, resulting in greater resolution for detecting genetic interactions (Supplementary Note 3). Despite these differences, we observed reasonable agreement among reciprocal genetic interaction pairs (Fig. 3b) and this correlation increased when we restricted comparisons to more confident interactions (Fig. 3b). Notably, agreement between reciprocal interactions improved ~50–60% when we normalized experimental factors responsible for systematic colony-size variation (Supplementary Fig. 3).

Finally, we compared SGA scores for 239 genetic interactions⁷ that overlapped with a smaller-scale, high-resolution liquid growth study⁵. We observed substantial quantitative agreement between genetic interactions identified in both studies (Fig. 3c). This was not simply attributable to extreme interactions because we also observed high correlation even for modest genetic interactions ($r = 0.73$; SGA score absolute value, $|\epsilon| < 0.4$). We also found high agreement when we compared our identified interactions to a set of manually confirmed synthetic lethal or sick interactions⁹ (Supplementary Fig. 4) as well as to a larger set of binary genetic interactions in the Biological General Repository for Interaction Datasets (BioGRID)¹¹ (Supplementary Note 4 and Supplementary Table 3).

Functional impact of systematic effects

We scored ~1,700 genome-wide SGA screens^{1,7} (Supplementary Note 1) to determine the effects of systematic experimental variation on reliable identification of fitness-based genetic interactions. We evaluated the functional utility of resultant positive and negative genetic interactions against common functional benchmarks: annotation to the same Gene Ontology (GO) biological process term or protein-protein interactions (Fig. 4a and Supplementary Fig. 5a). Although genetic interactions are not always expected to connect functionally related genes, they tend to be enriched among genes in the same biological process, thus annotation to the same GO terms serves as an objective

metric for quantifying the functional utility of a given dataset¹². Both negative and positive SGA score interactions overlapped with co-annotated gene pairs more frequently than interactions derived from a version of the SGA score without the normalization methods applied (Fig. 4a). At 30% precision, negative SGA scores recovered over fourfold more co-annotated gene pairs (Fig. 4a), suggesting that experimental variability contributes substantially to the false positive rate of large-scale SGA screens. The SGA score had similar improvement over genome-wide screens processed using the S score³, which does not account for many of the systematic effects we identified (Fig. 4a, Supplementary Fig. 5a and Online Methods). Despite capturing the most extreme genetic interactions, the S score appears to call false positives when applied to whole-genome screens because it does not account for batch effects (Supplementary Figs. 5c and 6 and Supplementary Table 1).

Accounting for systematic error in genome-wide genetic-interaction profiles yielded several-fold more functionally informative profiles than those obtained from nonnormalized data (~3,000 functionally related gene pairs at 30% precision using SGA score-derived profiles versus <10 functionally related gene pairs using nonnormalized colony size measurements; Fig. 4b). As observed for individual interactions, the batch effect was the main contributor to improved genetic-interaction profiles (Fig. 4b and Supplementary Figs. 5c and 6).

Relating genetic interactions to protein complexes

We reexamined the relationship between quantitative genetic interactions identified using the SGA score and physical association data for proteins. We focused on genetic interactions among 161 annotated protein complexes (Supplementary Note 5) for which more than one protein pair had been screened for genetic interactions⁷ and measured the frequency of positive and negative interactions in each complex (Fig. 5a, Supplementary Fig. 7 and Supplementary Data 2). Consistent with smaller-scale studies¹³, a large portion (92/161) of these complexes were significantly enriched for genetic interactions ($P < 0.05$, hypergeometric test). In the enriched complexes, the majority of genes were linked to one another either by pure positive (46%) or pure negative (37%) genetic interactions (Fig. 5a), confirming previous theoretical observations¹⁴.

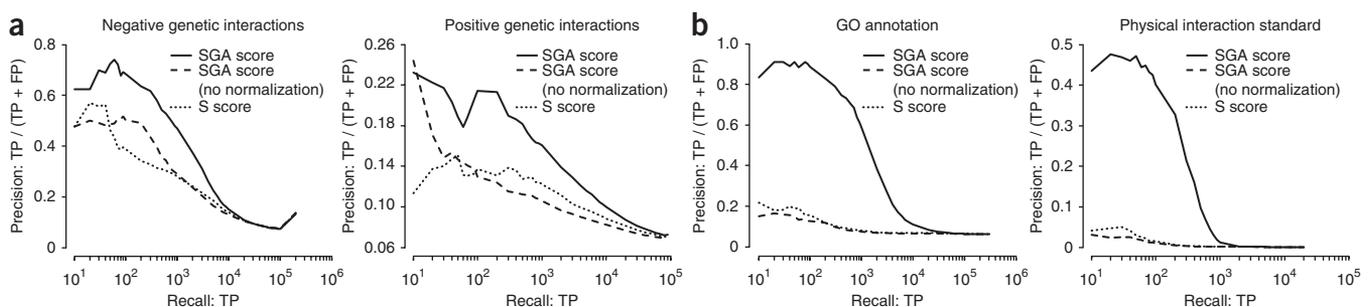
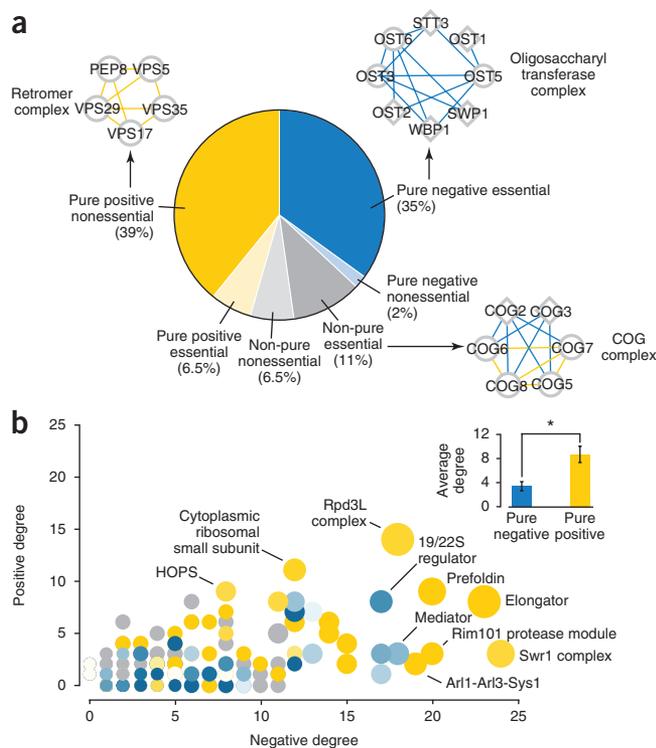


Figure 4 | Evaluation of functional information derived from genetic interactions. **(a)** Plots of precision versus recall (number of true positives (TP)) for negative and positive genetic interactions, as determined by the SGA score or the S score. An SGA score without normalization methods applied is also plotted. True positive interactions were defined as those involving both genes annotated to the same GO gold standard set of terms¹². Precision and recall were calculated as described previously¹². FP, false positive. **(b)** Plots of precision versus recall (number of TP) for genetic interaction profile similarities computed using the SGA score or the S score. An SGA score without normalization methods applied is also plotted. Pearson correlation was used to compute profile similarity for every pair of array mutant strains across profiles consisting of interactions with the 1,712 query mutant strains. True positive pairs were those for which both genes were annotated to the same GO gold standard set of terms (GO annotation)¹² or pairs encoding physically interacting proteins (physical interaction standard). Precision and recall were calculated as described previously¹².



Virtually all (94%) complexes characterized by negative genetic interactions had at least one essential gene (Fig. 5a), which is much higher than appreciated previously¹³. This observation suggests that essential complexes may contain internal redundancy, allowing the cell to tolerate loss of a single nonessential component of the complex (most interactions in our dataset were between nonessential genes), whereas additional perturbation often resulted in loss of complex function and impaired cell growth. In contrast, only a relatively small fraction of all protein complexes associated with only positively interacting gene pairs contained an essential gene (Fig. 5a). Complexes with mixed interactions (both positive and negative; 17.5%) also tended to be essential. For example, the highly organized conserved oligomeric Golgi (COG) complex, consisting of an essential and nonessential lobe¹⁵ had negative and positive intra-complex interactions, respectively. This demonstrates the resolution of the SGA score for capturing interactions that previously had been identified only by high-resolution competitive-growth assays⁴.

In addition to essentiality, the presence of a direct physical interaction identified in a yeast two-hybrid assay also influenced the type of genetic interaction observed in a protein complex. Proteins in a complex that physically interacted directly were nearly threefold more likely to have positive genetic interactions whereas complex members with no evidence of a direct physical interaction showed a modest preference toward negative genetic interactions (Supplementary Fig. 7d). The predominant type of genetic interaction we observed in a protein complex was predictive of its frequency of interaction with genes in other complexes (Supplementary Note 5). Nonessential genes in complexes connected by positive interactions had an average of twofold more genetic interactions compared to nonessential genes within complexes connected by negative genetic interactions (Fig. 5b). The fewer genetic interactions for essential complexes may reflect the inherent redundancy within these complexes; a single perturbation

is less likely to compromise complex activity, and thus, these nonessential genes exhibit fewer genetic interactions. Conversely, deletion of a nonessential gene in a positively connected complex may have a more severe impact on complex activity and, consequently, exhibited relatively more genetic interactions with the rest of the genome.

Figure 5 | Analysis of genetic interactions within and between protein complexes. (a) For 92 complexes enriched for negative and/or positive genetic interactions that we assembled, nonessential genes are represented as circles and essential genes as diamonds. Complexes connected by purely positive or negative genetic interactions are indicated by yellow and blue, respectively; gray denotes complexes connected by a mixture of positive and negative genetic interactions. (b) Degree analysis of the complex-complex genetic interaction network in which node color reflects the prevalence of positive (yellow) or negative (blue) genetic interactions in a complex. Gray nodes denote complexes for which too few gene pairs were screened to assess within-complex interactions. Node size indicates the number of proteins associated with the complex. Positive and negative degree are the number of positive and negative genetic interactions, respectively, for a given complex. Inset, number of between-complex interactions was measured for complexes connected by purely negative and purely positive genetic interactions. Error bars, s.e.m. ($n = 37$, positive and $n = 25$, negative; $*P < 0.002$ by a rank-sum test).

is less likely to compromise complex activity, and thus, these nonessential genes exhibit fewer genetic interactions. Conversely, deletion of a nonessential gene in a positively connected complex may have a more severe impact on complex activity and, consequently, exhibited relatively more genetic interactions with the rest of the genome.

Defining positive interaction subclasses

Even though both positive and negative genetic interactions are enriched in physical complexes, the large majority of genetic interactions did not overlap with physical interactions from high-throughput assays^{16–19}. We found 1,182 pairs of protein complexes that showed a substantial enrichment for positive interactions connecting them (false discovery rate of 5%; Supplementary Note 5). The numerous positive genetic interactions that do not overlap with protein-protein interactions may reflect functional rather than physical relationships between complexes (Fig. 6a).

Single- and double-mutant phenotypes have been compared in small-scale studies to identify specific classes of positive interactions, including genetic suppression^{5,8}. We used the SGA score to examine the potential for genetic suppression between protein complexes on a global scale (Supplementary Note 6 and Supplementary Data 3). We observed a surprising number of suppression interactions across complex pairs, which suggest instances in which loss-of-function mutations in one complex rescue growth defects associated with loss-of-function mutations in a second complex. We constructed a network providing a global view of suppression interactions between protein complexes (Fig. 6a).

In addition to a cross-complex suppression interaction identified previously in a high-resolution growth competition assay⁴, we validated several new suppression interactions in our network (Fig. 6a and Supplementary Fig. 8). We confirmed loss-of-function suppression interactions involving the Rim101 signaling pathway genes *RIM8*, *RIM9* or *DFG16*, multivesicular body sorting proteins encoded by *DID4* and *VPS24*, which comprise an ESCRT-III subcomplex, and the AAA-type ATPase gene, *VPS4* (Fig. 6a,b and Supplementary Fig. 9a,b). A functional relationship between Rim101 signaling and multivesicular body sorting has been established previously^{20,21}. *RIM101* encodes a transcription factor activated in response to alkaline growth conditions

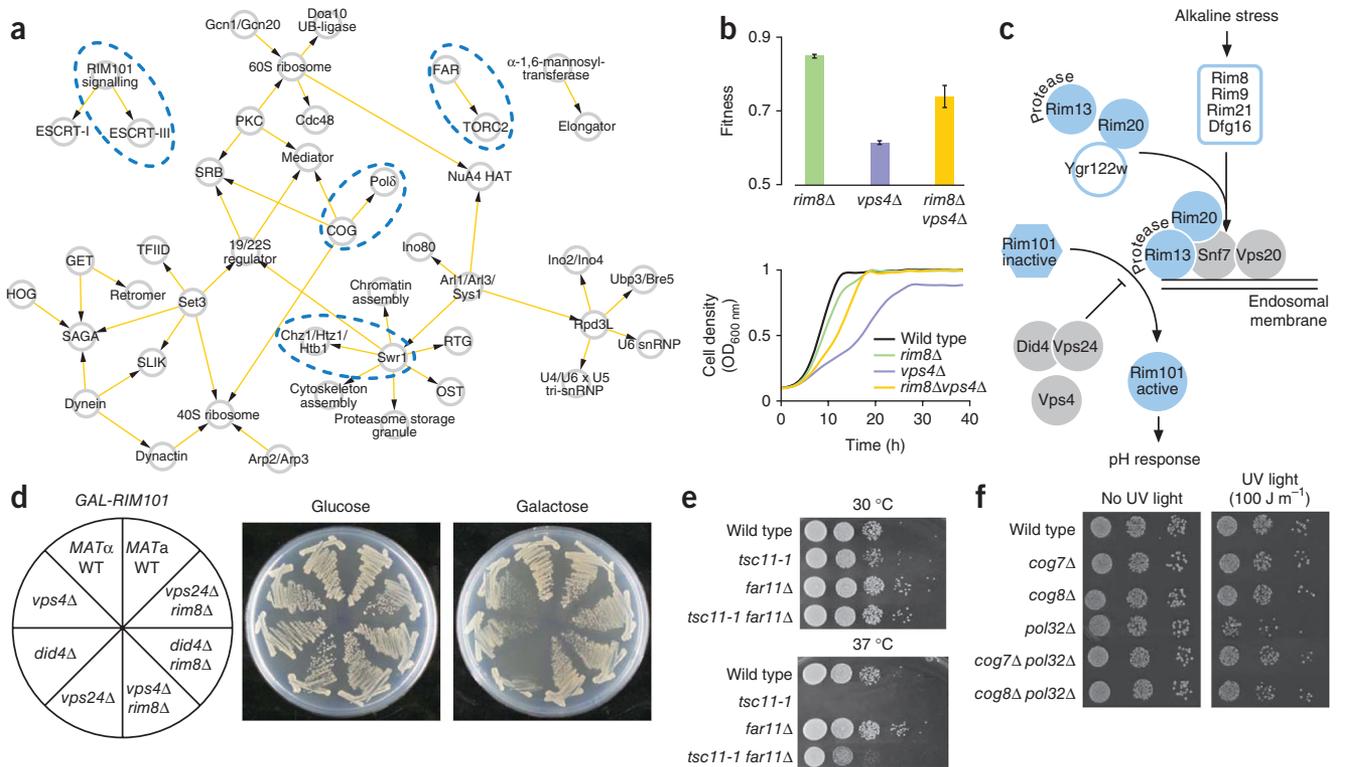


Figure 6 | Cross-complex genetic suppression network revealed by quantitative genetic interaction analysis. **(a)** A network illustrating suppression interactions between protein complexes (nodes). Edges indicate positive SGA interactions classified as suppression; arrows point to the complex whose fitness defect was suppressed. **(b)** Colony size-derived single- and double-mutant fitness plotted for the indicated strains (top). Error bars for single mutants, s.e.m. derived from bootstrapping ($n=800$); error bars for double mutant, s.d. ($n=4$). Liquid growth profiling⁵ of the same strains (bottom). **(c)** Schematic showing activation of the Rim101 pathway in response to alkaline stress. **(d)** Growth of the indicated strains on glucose and galactose, with plated strains indicated (left). WT, wild type. **(e, f)** Serial dilution growth assays of the indicated yeast strains at the indicated temperatures **(e)** and after exposure to UV light.

via proteolytic cleavage at endosomal membranes (Fig. 6c). Our suppression network highlighted that deletion of genes encoding upstream signaling components of the Rim101 pathway suppressed fitness defects associated with deletion of *DID4*, *VPS24* or *VPS4*; this may occur because a defect in upstream signaling prevents constitutive activation caused by loss of function of downstream negative regulators. We confirmed that *did4Δ*, *vps24Δ* and *vps4Δ* mutants were sensitive to *RIM101* overexpression, and this sensitivity was rescued by deletion of *RIM8* (Fig. 6d), a putative upstream component of the Rim101 pathway²¹ (Fig. 6c). Under conditions in which the Rim101 pathway is activated and required for viability, in the presence of lithium chloride (LiCl), suppression is observed in the opposite direction²¹. These observations highlight the importance of considering condition-specificity when inferring pathway architecture based on genetic interactions⁶. In addition to characterized pathway components, our genetic interaction analysis placed *DFG16*, an uncharacterized gene associated with Rim101 signaling²², upstream in the pathway along with *RIM8* and *RIM9*. Another gene previously implicated in the Rim101 signaling, *YGR122W*²³, likely functions further downstream and closer to the Rim13 protease (Fig. 6c and Supplementary Fig. 9a–c).

Disruption of the FAR complex, originally implicated in cell-cycle control²⁴, rescued growth defects associated with TORC2 kinase complex mutant alleles, *tor2-29* and *tsc11-1* (Fig. 6a,e and Supplementary Fig. 10a,b). Moreover, *FAR11* deletion suppressed

actin polarization defects of a *tsc11-1* mutant in a nonpermissive condition (37°C; Supplementary Fig. 10c). These results suggest that the FAR complex may function downstream to negatively regulate TORC2 function in actin organization. Similar to TORC2, FAR complex members are conserved from yeast to humans, and mammalian Far protein orthologs belong to a multiprotein complex that contains the PP2A phosphatase²⁵. Suppression of TORC2 growth and actin polarity defects was also achieved by loss of *PPG1* (Supplementary Fig. 10a–c), which encodes a PP2A-related phosphatase²⁶. Thus, it is possible that the FAR complex mediates its function by working with Ppg1 to dephosphorylate and inactivate proteins that normally control actin-based cell polarity.

Genes involved in chromatin and secretory functions act as hubs in the global genetic interaction map⁷ and our suppression network exhibited a similar topology (Fig. 6a). Specifically, protein complexes involved in chromatin modification or secretion suppressed growth defects associated with disruption of several different pathways and/or complexes. One such interaction involved suppression of DNA polymerase delta (*Polδ*) mutants by COG complex disruption (Fig. 6a and Supplementary Fig. 11). Unlike *Polδ*, which functions in the nucleus, the COG complex is important for establishment and maintenance of Golgi apparatus structure and function¹⁵. To explore this genetic relationship, we investigated whether *cogΔ* deletion mutants also rescued the UV light-sensitivity of a strain lacking a nonessential *Polδ* gene, *POL32* (Fig. 6f). The *cog7Δpol32Δ* and *cog8Δpol32Δ* double

mutants grew as well as *cog7Δ* and *cog8Δ* strains after UV-light exposure, confirming the COG-Polδ suppression interactions and the ability of positive interactions to capture broad connections between functionally diverse genes.

DISCUSSION

Several studies have highlighted the importance of quantitative fitness analysis for dissecting pathways and complexes^{4,5,8,27}. Although our primary goal was to develop a scoring system for extracting accurate fitness estimates and genetic interactions from genome-scale screens in budding yeast, our methodology can be applied to map genetic interactions in other organisms, including different yeast species^{28,29} and bacteria³⁰, for which colony-based assays for genetic interactions exist. Furthermore, our methodology can be extended beyond genetic interactions to different array-based interaction mapping technologies such as yeast two-hybrid¹⁶ or protein-fragment complementation assay¹⁷, which rely on colony growth to identify genes encoding physically interacting proteins.

Normalization of experimental factors affecting colony growth and experimental grouping of screens was key to extracting fitness-based genetic interactions. Because SGA experiments are manipulated robotically, these experimental factors were surprisingly reproducible and, thus, amenable to estimation and normalization, which enhanced the measurement precision without sacrificing throughput. Not surprisingly, given its impact in other genome-wide studies³¹, batch effect presented a major challenge. Proper experimental design, including randomization of query gene screening order, especially with respect to their functional roles, was critical to distinguish between biological and experimental factors unique to screens belonging to a common batch. Batch effect correction may present a greater challenge for smaller-scale and functionally biased studies in which true genetic interactions and colony variance owing to batch effects are not easily distinguished. Repeating screens independently or pairing them with control screens for detailed comparison may highlight batch-specific effects. Furthermore, randomizing positions of array strain replicates may also minimize other sources of experimental error associated with plate, spatial and competition effects.

Our global genetic interaction map⁷ revealed that, contrary to previous understanding, both negative and positive genetic interactions tend to occur between, rather than within, complexes and pathways. Moreover, suppression interactions observed between complexes in distinct subcellular compartments highlighted the potential for genetic interactions to uncover important and unexpected functional relationships. Consistent with our findings, an *in silico* study found that negative interactions occur more frequently between genes with overlapping function whereas positive interactions are observed between functionally distinct metabolic pathways³². However, unlike our genome-wide survey indicating an increased prevalence of negative genetic interactions⁷, this theoretical analysis suggests that positive interactions are surprisingly more abundant than negative interactions in *S. cerevisiae* and *Escherichia coli* metabolic networks³². Whether this is a specific characteristic of metabolic networks or a more global network property related to gene essentiality remains to be explored.

Genome-wide application of our method is an important step toward moving beyond abstract gene function predictions toward

a specific, mechanistic wiring diagram of the cell. Continued quantitative analysis of SGA experiments should enable measurement of all possible yeast double-mutant combinations, and our analysis suggests the resolution of SGA scored interactions may, in some cases, even be sufficient for inferring pathway architecture (Fig. 6b,d). Furthermore, strategies for normalizing experimental factors will become more critical as we expand genetic interaction maps to different conditions, organisms and complex relationships involving more than two genes. This enormous interaction space emphasizes the need for highly scalable approaches and tools capable of high-resolution measurement of single and combined mutant phenotypes.

METHODS

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

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

ACKNOWLEDGMENTS

We thank R. Mani, F. Roth, B. Papp, T. Maeda and T. Hays for helpful discussions and critical comments. This work was supported by Genome Canada through the Ontario Genomics Institute (2004-OGI-3-01), the Canadian Institutes of Health Research (GSP-415-67; C.B. and B.A.) (MOP-79368; G.W.B.), the US National Institutes of Health (1R01HG005084-01A1; C.L.M., Y.K. and S.B.) and the National Science Foundation (DBI 0953881, MCB 0918908; C.L.M. and Y.K.) and partially supported by a seed grant from the Minnesota Supercomputing Institute (C.L.M. and Y.K.).

AUTHOR CONTRIBUTIONS

M.C., C.L.M., C.B. and B.A. conceived and coordinated the project. C.L.M. and A.B. designed and implemented the algorithm. A.B., C.L.M., Y.K., J.K. and S.B. performed statistical analysis. J.-Y.Y., B.-J.S.L., J.O., G.W.B. and M.C. validated experiments. H.D. and K.T. analyzed and processed images. M.H., D.H., G.D.B. and O.G.T. provided statistical insight. A.-C.G. provided biological insight. M.C., C.L.M., C.B. and A.B. prepared the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturemethods/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

1. Baryshnikova, A. *et al.* Synthetic genetic array (SGA) analysis in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Methods Enzymol.* **470**, 146–180 (2010).
2. Dixon, S.J., Costanzo, M., Baryshnikova, A., Andrews, B. & Boone, C. Systematic mapping of genetic interaction networks. *Annu. Rev. Genet.* **470**, 145–179 (2009).
3. 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).
4. Breslow, D.K. *et al.* A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat. Methods* **5**, 711–718 (2008).
5. St Onge, R.P. *et al.* Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat. Genet.* **39**, 199–206 (2007).
6. Avery, L. & Wasserman, S. Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet.* **8**, 312–316 (1992).
7. Costanzo, M. *et al.* The genetic landscape of a cell. *Science* **327**, 425–431 (2010).
8. Drees, B.L. *et al.* Derivation of genetic interaction networks from quantitative phenotype data. *Genome Biol.* **6**, R38 (2005).
9. Tong, A.H. *et al.* Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813 (2004).
10. Deuschbauer, A.M. *et al.* Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925 (2005).

11. Breitkreutz, B.J. *et al.* The BioGRID interaction database: 2008 update. *Nucleic Acids Res.* **36**, D637–D640 (2008).
12. 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).
13. Bandyopadhyay, S., Kelley, R., Krogan, N.J. & Ideker, T. Functional maps of protein complexes from quantitative genetic interaction data. *PLoS Comput. Biol.* **4**, e1000065 (2008).
14. Segre, D., Deluna, A., Church, G.M. & Kishony, R. Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83 (2005).
15. Ungar, D., Oka, T., Krieger, M. & Hughson, F.M. Retrograde transport on the COG railway. *Trends Cell Biol.* **16**, 113–120 (2006).
16. Yu, H. *et al.* High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104–110 (2008).
17. Tarassov, K. *et al.* An in vivo map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008).
18. Gavin, A.C. *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636 (2006).
19. Krogan, N.J. *et al.* Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637–643 (2006).
20. Mitchell, A.P.A. VAST staging area for regulatory proteins. *Proc. Natl. Acad. Sci. USA* **105**, 7111–7112 (2008).
21. Hayashi, M., Fukuzawa, T., Sorimachi, H. & Maeda, T. Constitutive activation of the pH-responsive Rim101 pathway in yeast mutants defective in late steps of the MVB/ESCRT pathway. *Mol. Cell. Biol.* **25**, 9478–9490 (2005).
22. Barwell, K.J., Boysen, J.H., Xu, W. & Mitchell, A.P. Relationship of DFG16 to the Rim101p pH response pathway in *Saccharomyces cerevisiae* and *Candida albicans*. *Eukaryot. Cell* **4**, 890–899 (2005).
23. Rothfels, K. *et al.* Components of the ESCRT pathway, DFG16, and YGR122w are required for Rim101 to act as a corepressor with Nrg1 at the negative regulatory element of the DIT1 gene of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **25**, 6772–6788 (2005).
24. Kemp, H.A. & Sprague, G.F. Jr. Far3 and five interacting proteins prevent premature recovery from pheromone arrest in the budding yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **23**, 1750–1763 (2003).
25. Goudreault, M. *et al.* A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol. Cell. Proteomics* **8**, 157–171 (2009).
26. Posas, F. *et al.* The gene PPG encodes a novel yeast protein phosphatase involved in glycogen accumulation. *J. Biol. Chem.* **268**, 1349–1354 (1993).
27. Van Driessche, N. *et al.* Epistasis analysis with global transcriptional phenotypes. *Nat. Genet.* **37**, 471–477 (2005).
28. Dixon, S.J. *et al.* Significant conservation of synthetic lethal genetic interaction networks between distantly related eukaryotes. *Proc. Natl. Acad. Sci. USA* **105**, 16653–16658 (2008).
29. Roguev, A. *et al.* Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* **322**, 405–410 (2008).
30. Babu, M. *et al.* Systems-level approaches for identifying and analyzing genetic interaction networks in *Escherichia coli* and extensions to other prokaryotes. *Mol. Biosyst.* **5**, 1439–1455 (2009).
31. Leek, J.T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* **11**, 733–739 (2010).
32. He, X., Qian, W., Wang, Z., Li, Y. & Zhang, J. Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. *Nat. Genet.* **42**, 272–276 (2010).
33. Hughes, T.R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126 (2000).
34. Jasnos, L. & Korona, R. Epistatic buffering of fitness loss in yeast double deletion strains. *Nat. Genet.* **39**, 550–554 (2007).
35. Warringer, J., Ericson, E., Fernandez, L., Nerman, O. & Blomberg, A. High-resolution yeast phenomics resolves different physiological features in the saline response. *Proc. Natl. Acad. Sci. USA* **100**, 15724–15729 (2003).
36. Brauer, M.J. *et al.* Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol. Biol. Cell* **19**, 352–367 (2008).
37. Hartman, J.L.t. & Tippery, N.P. Systematic quantification of gene interactions by phenotypic array analysis. *Genome Biol.* **5**, R49 (2004).



ONLINE METHODS

SGA genetic interaction score. The source code for normalization procedures is available in **Supplementary Software**. The raw colony size data are available at <http://csbio.cs.umn.edu/SGAScore/>. Details of the normalization methods and genetic interaction score are described in **Supplementary Note 1**.

Strains. All deletion mutant strains used for serial dilution, actin staining and RIM101 expression experiments were derivatives of BY4741 or Y7092, the construction which has been described previously¹. Two different temperature-sensitive alleles were used in this study: *TSQ658* (*MAT α tsc11-1::natMX can1 Δ ::STE2pr-Sp_his5*; *lyp1 Δ ::STE3pr-LEU2*; *his3 Δ 1 leu2 Δ 0 ura3 Δ 0 LYS2*) and *TSQ838* (*MAT α tor2-29::natMX can1 Δ ::STE2pr-Sp_his5*; *lyp1 Δ ::STE3pr-LEU2*; *his3 Δ 1 leu2 Δ 0 ura3 Δ 0 LYS2*). All double-mutant strains were constructed by SGA.

SGA screens. We performed 1,712 genome-wide SGA screens using eight BioMatrix Colony Arrayer Robots (S&P Robotics). A single SGA screen takes approximately 3 weeks, and the entire dataset was collected over 2 years, as described previously^{1,7}. Double-mutant SGA plates were digitally photographed, and colony areas were obtained from the images using the Colony software (S&P Robotics).

Spot dilutions: *tsc11-1* suppression. Overnight cultures were serially diluted 20-fold and spotted onto agar medium as indicated. Strains were grown for 2 d at 30 °C or 37 °C as indicated.

Actin staining. Cultures were grown at 30 °C to early log phase ($OD_{600\text{ nm}}$ of ~0.1–0.2) and then shifted to nonpermissive

temperature (37 °C) for 6 h. Actin staining was done as described previously³⁸ without added fluorescent brightener. Stained cells were imaged using a DMI 6000B fluorescence microscope (Leica Microsystems) equipped with a spinning-disk head, an argon laser (458 nm, 488 nm and 514 nm; Quorum Technologies) and Imagem charge-coupled device camera (C9100-13; Hamamatsu). Cells were imaged in seven z stacks with 0.3- μm intervals, and all z stacks were collapsed into one extended focus.

UV-light sensitivity: *pol32 Δ* suppression. UV-light sensitivity assay was performed as described previously³⁹.

S score comparative analysis. Comparative analyses were conducted by processing raw colony size data, derived from SGA, using the S-score method that was reimplemented as described previously³.

Protein-protein interaction standard. Protein-protein interaction data were downloaded from BioGRID⁴⁰ on 17 September 2009.

Comparison of single-mutant fitness measurements. Fitness or growth-rate data were downloaded from the supplementary websites of the respective publications and used directly for all studies except for reference 35, which was exponentially transformed to reflect mutant fitness.

38. Friesen, H. *et al.* Characterization of the yeast amphiphysins Rvs161p and Rvs167p reveals roles for the Rvs heterodimer in vivo. *Mol. Biol. Cell* **17**, 1306–1321 (2006).
39. Bellaoui, M. *et al.* Elg1 forms an alternative RFC complex important for DNA replication and genome integrity. *EMBO J.* **22**, 4304–4313 (2003).
40. Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).