Supporting Information

Dittmar et al. 10.1073/pnas.1219582110

SI Materials and Methods

Yeast Strains and Media. Standard yeast growth media and strain-manipulation techniques were used throughout this work (1). Media were supplemented with 300 μg/mL G418 (Mediatech) for selection of the KanMX marker and 100 μg/mL clonNAT (Werner BioAgents) for selection of the NatMX marker. All yeast strains for screens and verification experiments were from the *MAT***a** and *MAT*α gene disruption libraries (2). Several gene disruptions were transferred from the gene-disruption library strains to a *RAD5* derivative of W303 by PCR amplification of the disruption, including ≥300 bp of flanking homologous DNA, followed by lithium acetate-mediated transformation of the new strain as described previously (3). All newly made gene disruptions were backcrossed at least once to a wild-type strain to ensure 2:2 segregation of the transferred marker.

CLIK Software. CLIK (Cutoff Linked to Interaction Knowledge) is available as open-source software (GNU General Public License) through the laboratory of R.R. (www.rothsteinlab.com/tools/apps/clik) and through SourceForge (https://sourceforge.net/p/clik). Here, users may run CLIK as a Web application and download the source code. For inputted data, CLIK scatter plots are created by placing the rank-order list of genes from a single screen on the *x* axis and the same rank order on the *y* axis. A point is plotted for each interaction between ORFs on the axes (Fig. 1 *A* and *B*). Once plotted, the density of each point on the graph is calculated by centering a virtual square bin over each point and dividing the number of points that fall within the bounds of the bin by the area. By default, bin sizes are automatically calculated by moving a window of various sizes across the rank list entered into CLIK. For each window, a score is calculated by dividing the number of interactions within the window by the window size. The window size with the greatest score is then used as the *x* and *y* dimension of the virtual bin that CLIK uses to calculate point density. Manual entry of bin sizes is also allowed to aid in the comparison of multiple screens (to normalize across analyses).

To highlight significant density values on CLIK graphs, a threshold density value is determined. This threshold is comprised of two parts. The first is the background level of interaction density, referred to as the "background density" (δ_b), which is inherent in each interactome being analyzed. To calculate δ_b, a control dataset is first generated by randomizing the rank-order positions of the original ORFs. The distribution of densities from this random set approximates a normal distribution. The density corresponding to the 95th percentile is set as δ_b. However, this value does not completely eliminate nonsignificant density values within the ordered dataset because it simply represents background noise attributable to random variance. Therefore, an additional component of the threshold is needed. For this second component, we use a value of 0.5 SD from the mean of the ordered dataset density values. Any density value greater than the sum of these two components is deemed significant. The determination of significant density values is summarized in Eq. **S1**:

$$\delta_{sig} = \{x \in \delta_o | x > 0.5\sigma_o + \mu_o + \delta_b\}, \quad \textbf{[S1]}$$

where δ_{sig} is the set of significant densities for the ordered dataset, δ_o is the set of densities from the ordered dataset, μ_o is the mean density of the ordered dataset, and σ_o is the SD of the ordered dataset.

The value 0.5 σ_o was chosen after varying this value from 0.0 to 1.0 and seeing that varying these values in the range of 0.2–0.8 did not significantly affect screen cutoff (Fig. S7*A*). Thus, the midpoint value of 0.5 was chosen.

The threshold value calculated by Eq. **S1** ensures that only density values significantly above background are considered. The threshold's effect on determining significant densities can be seen by analyzing the results from the *DMC1* and *TOP1* mutant allele (*top1-T_{722}A*) screens. Poorly performing screens, such as the *DMC1* synthetic dosage-lethality (SDL) screen (Fig. S1), do not significantly organize interacting genes at the top of the rank order and, thus, do not contain high interaction-density values. Conversely, screens that organize interacting genes at the top of the rank order, such as *top1-T_{722}A*, form strong CLIK groups containing high interaction values (Fig. 2*A*). As a result, the distribution of densities from poorly performing screens is much closer to a random distribution than a well-performing screen with a strong CLIK group (Fig. S7*B*). Thus, the threshold significant value for poorly performing screens is more stringent than for screens that perform well. For example, the density threshold value for the *DMC1* screen is ∼3 σ from the mean, whereas the threshold for the *top1-T_{722}A* screen is ∼1.645 σ from the mean.

To visualize significant interaction densities, all plot points within δ_{sig} are colored according to a scale; every other point is colored gray. The scale consists of 12 colors in a gradient (see scale in Fig. 1). Light green encompasses the minimum density threshold value plus 1 σ_o. Subsequent colors on the scale are 1 σ_o apart. However, this range may be too wide for some screens with high interaction densities and high σ_o (e.g., the cisplatin and *top1-t_{722}A* screens in Fig. 2). Therefore, if the maximum density is at least twice that of the minimum density in δ_{sig}, then the color scale is rescaled so that black represents the maximum significant density. In this case, the 10 other scale colors are evenly distributed between the minimum and maximum significant density values. The minimum and maximum densities on the color scale may also be manually set by the user through the Web interface.

CLIK produces an image of the CLIK graph and the details of any CLIK groups present within the graph are listed, such as the ORFs, the mean and maximum interaction densities, etc. The text of the individual ORFs is colored according to the corresponding interaction density on the diagonal of the CLIK graph to easily ascertain the CLIK determined cutoff position (i.e., transitions from color to gray).

Additional CLIK Options. There are several parameters that may be adjusted through the Web interface before performing CLIK analysis; these are as follows.

Interaction reciprocality. The density on the CLIK graphs tends to be mirrored across the diagonal of the graph; however, directionality may be observed because, in some cases, an interaction between a bait and prey has been annotated, but the converse has not. Selecting the "Consider all Interactions Reciprocal" option before running CLIK will eliminate this directionality by interpreting all annotated interactions as bidirectional. Conversely, the user may wish to only consider interactions that have been observed in both directions (i.e., A → B and B → A). To do this, users may select the "'Only Consider Reciprocal Interactions" option before running CLIK.

Noise reduction. On some CLIK graphs, vertical and horizontal lines appear because of promiscuously interacting ORFs included in the analysis. To reduce the influence of these ORFs on CLIK, users may adjust the "Noise Reduction" parameter option before submitting their rank-order list for analysis. This drop-down menu allows users to omit ORFs with more than "X"

number of interactions within the list they submit for CLIK analysis. The default value of "X" at the time of publication is 400 interactions.

Interactions to consider. To discern the influence of physical interactions (e.g., two hybrid, FRET, cofractionation, etc.) from genetic ones (e.g., dosage lethality, snthetic lethality, positive genetic, etc.) on CLIK graphs, we added an "Interactions to Consider" section on the Web interface. All interaction types within the BioGRID database are listed and selected by default; however, users may modify the settings in this section to consider/omit interaction types as they see fit before running CLIK.

SDL Screens. SDL screens were performed as described previously (4). The *top1-$T_{722}A$* set 2 screen data from ref. 4 was analyzed for this study. ORFs for *SPC110* and *DMC1* were PCR-amplified and cloned by recombination into plasmid pWJ1512 to make copper-inducible query genes for SDL screens. Gene expression was induced by pinning onto plates containing 100 μM $CuSO_4$. Verification of screen results was performed after choosing individual strains from the gene-disruption library and testing 16 replicate colonies for sensitivity to query gene expression using the selective ploidy ablation protocol (4) (Dataset S2). Strains were considered to be validated if the experimental strain average growth ratio differed by more than 4 SDs from the control population mean.

Cisplatin Sensitivity. Gene-disruption strains identified as members of the cisplatin-sensitive CLIK group, plus 100 additional gene disruptions beyond the CLIK-derived cutoff, were chosen from the gene-disruption library, so that cisplatin sensitivity could be tested with 16 replicate colonies. Sensitivity was determined after pinning the strains to plates containing 83 μM cisplatin and measuring growth compared with multiple *his3*Δ control sets (Dataset S2-2).

Rapamycin-Sensitivity Screens. Gene-disruption strains were pinned from agar to 0.1 mL yeast extract peptone-dextrose (YPD) plus G418 (Mediatech) and were grown for 48 h at 30 °C to ensure all strains reached stationary phase. Cultures were then diluted 1:10 in water and quadruplicated onto YPD plus G418 plus rapamycin (Sigma-Aldrich) plates at a density of 1,536 colonies per plate. Plates were incubated at 30 °C, and images were captured using a flatbed scanner at ∼20, 24, 30, and 48 h. Colony sizes were quantified using ScreenMill software (5). The average colony sizes for rapamycin treatment at each time point were compared, and the incubation time at each concentration that gave the most similar absolute growth levels were chosen for analysis. Ranked lists based on growth ratios were used for CLIK analysis. For validation, the top most significant 400, 300, and 200 rapamycin-sensitive strains from the 4, 10, and 16 nM screens, respectively, were chosen, to ensure testing beyond the CLIK-derived cutoffs. Because of screen overlap, a total 624 strains were chosen from the disruption library, cultured, and spotted onto the drug plates, and each strain was spotted 16 times. Each plate also included multiple replicas of the *his3*Δ control strain.

Dilution assays in Fig. 5*A* were performed on 12 strains by inoculating 0.1 mL of YPD and incubating at 30 °C for 2 d. Tenfold serial dilutions were spotted onto YPD plus rapamycin plates using a Singer RoToR robotic workstation. Images were taken at ∼48 h.

Gene-Ontology Enrichment. Gene-ontology (GO) enrichment of the validated ORFs from the cisplatin screen and 4 and 16 nM rapamycin screens was calculated using GOrilla (6); only process terms were considered (Dataset S1-2, S1-3, and S1-4). In Fig. S4*A*, only "Gold-Standard" GO terms present in Dataset S1-3 and S1-4 were considered (7).

1. Sherman F, Fink GR, Hicks JB (1983) *Methods in Yeast Genetics* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).
2. Winzeler EA, et al. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901–906.
3. Schiestl RH, Gietz RD (1989) High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* 16(5-6):339–346.
4. Reid RJ, et al. (2011) Selective ploidy ablation, a high-throughput plasmid transfer protocol, identifies new genes affecting topoisomerase I-induced DNA damage. *Genome Res* 21(3):477–486.
5. Dittmar JC, Reid RJ, Rothstein R (2010) ScreenMill: A freely available software suite for growth measurement, analysis and visualization of high-throughput screen data. *BMC Bioinformatics* 11:353–363.
6. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z (2009) GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48–54.
7. Myers CL, Barrett DR, Hibbs MA, Huttenhower C, Troyanskaya OG (2006) Finding function: Evaluation methods for functional genomic data. *BMC Genomics* 7:187–201.

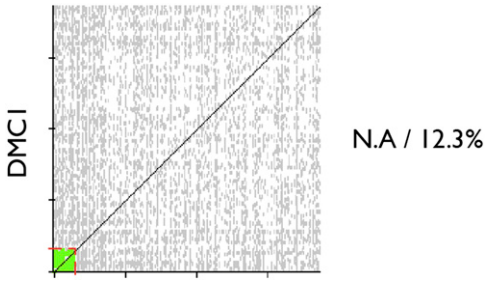

Fig. S1. CLIK graph and validation of the *DMC1* SDL screen. The CLIK graph was produced as in Fig. 2. Only 10 of the top 100 ORFs in the rank list validated, so a receiver operating characteristic curve was not constructed. Instead, validation rate with respect to rank order is shown in Fig. S2.

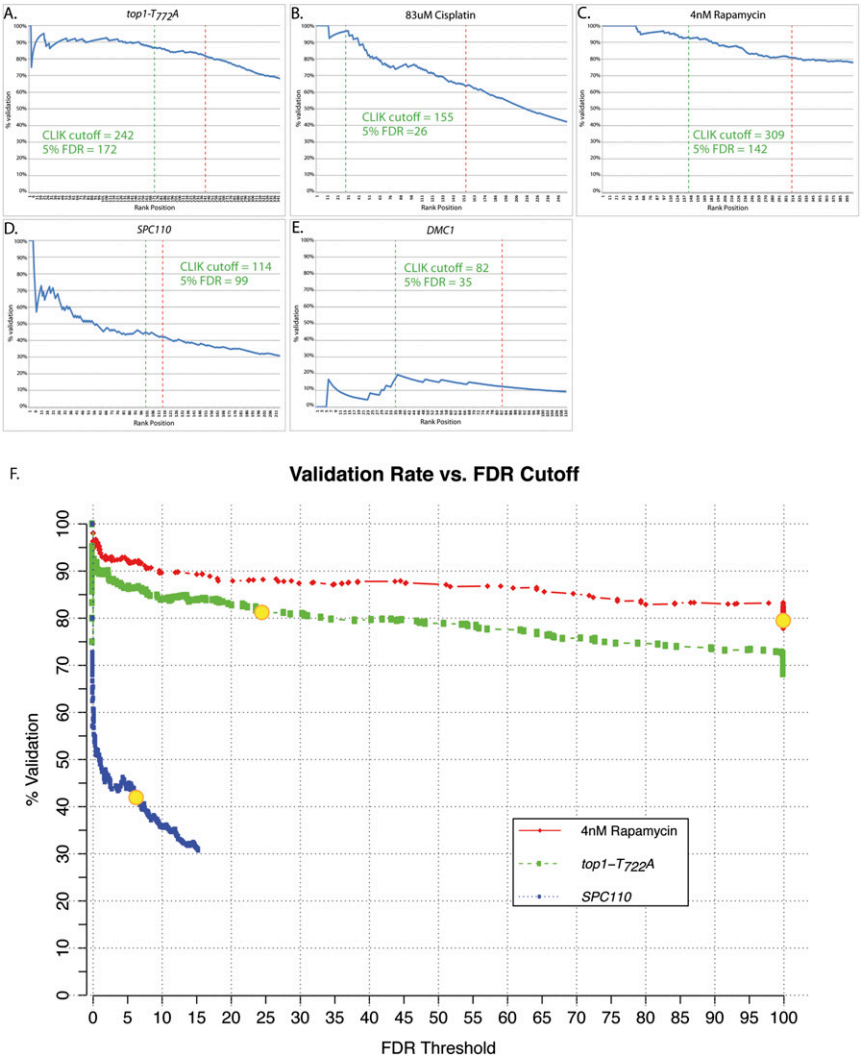

Fig. S2. (*A*–*E*) Validation rate vs. rank order. Percentage of validation was calculated at each rank (growing window) for all mutants tested [CLIK-derived cutoff plus the next 100 mutants (*A*–*D*) or plus the next 28 mutants (*E*)] to produce the graphs. Red vertical lines indicate the position in the rank order of the CLIK-derived cutoff, and green lines indicate the 5% FDR cutoff. (*A*) *top1-T*$_{722}$*A*–sensitivity screen, (*B*) Cisplatin-sensitivity screen. (*C*) Rapamycin-sensitivity screen. (*D*) *SPC110* SDL screen. (*E*) *DMC1* SDL screen. (*F*) FDR vs. validation plots. FDR values and the corresponding validation rates were calculated for the *top1-T*$_{722}$*A*–sensitivity screen, 4 nM rapamycin-sensitivity screen, and *SPC110* SDL screen. Yellow circles indicate the CLIK-derived cutoff for each screen.

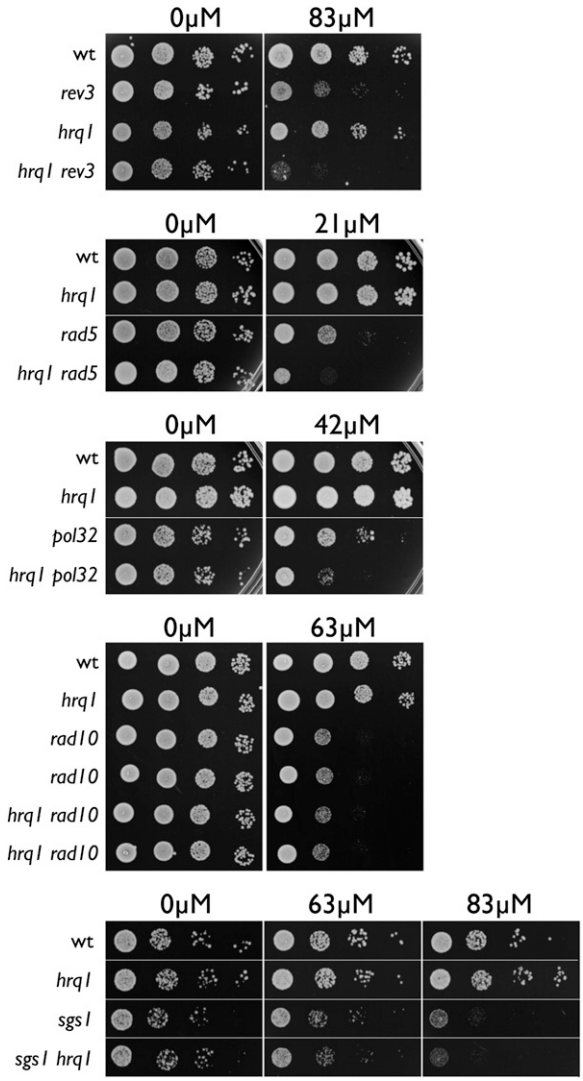

Fig. S3. Additional cisplatin-sensitivity spot assays. Spot assays were performed as in Fig. 3 and described in *SI Materials and Methods*. The *hrq1*-null mutant was crossed to *rev3*, *rad5*, *pol32*, *rad10*, and *sgs1* mutant strains to generate single and double mutant combinations as indicated. Cultures were serially diluted and spotted onto plates with and without 21, 42, 63, or 83 μM cisplatin to show drug sensitivity. *hrq1* shows a synergistic effect on cisplatin sensitivity when combined with *rev3*, *rad5*, or *pol32* but shows epistasis in combination with both *rad10* and *sgs1*.

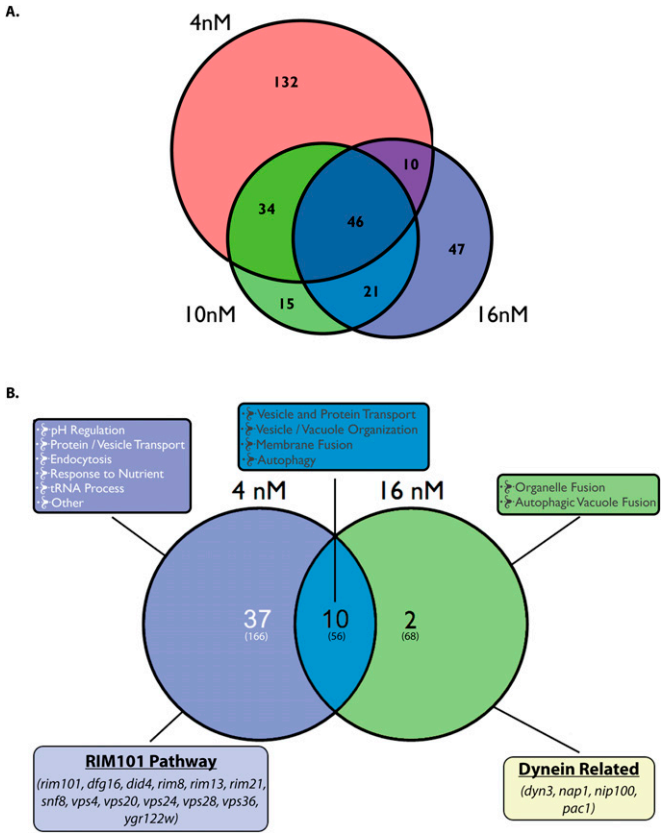

Fig. S4. (*A*) Venn diagram depicting the overlap between the validated strains within CLIK groups at each rapamycin concentration tested. Strains that did not grow on both the control, and experimental plates were not analyzed. Duplicates, if present, were only considered once in each set. Diagram generated with BioVenn (1). (*B*) Venn diagram depicting the GO process enrichment of validated hits from the 4 and 16 nM rapamycin-sensitivity screens present in Dataset S2. Diagram generated with Venny (2). Within the circles, large numbers represent the number of GO process terms shared between and exclusive to the 4 and 16 nM screens. The numbers in parenthesis indicates the number of genes comprising each group. At the bottom of the figure, validated genes in the RIM101 pathway (exclusive to 4 nM) and those related to dynein (exclusive to 16 nM) are indicated.

1. Hulsen T, de Vlieg J, Alkema W (2008) BioVenn—a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9:488.
2. Oliveros JC (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. Available at http://bioinfogp.cnb.csic.es/tools/venny/index.html.

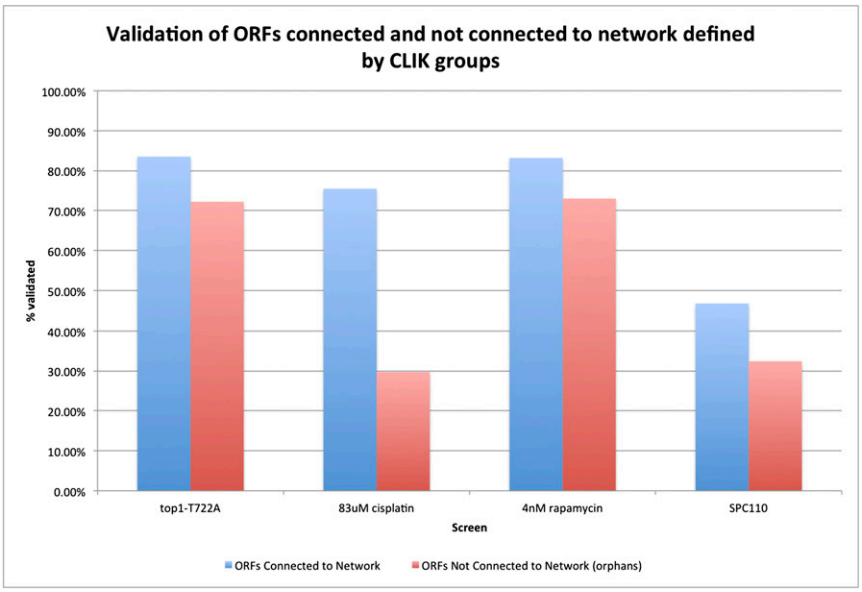

Fig. S5. Validation rate of ORFs within CLIK groups. ORFs within the CLIK groups of each screen in Fig. 2 were separated into two groups: those that had known interactions to other ORFs within the CLIK group and those that did not. The validation rate of each of these groups is shown for the four screens that contained well-defined CLIK groups at the top of the rank order. Strains that did not grow on both the control and experimental plates were not considered.

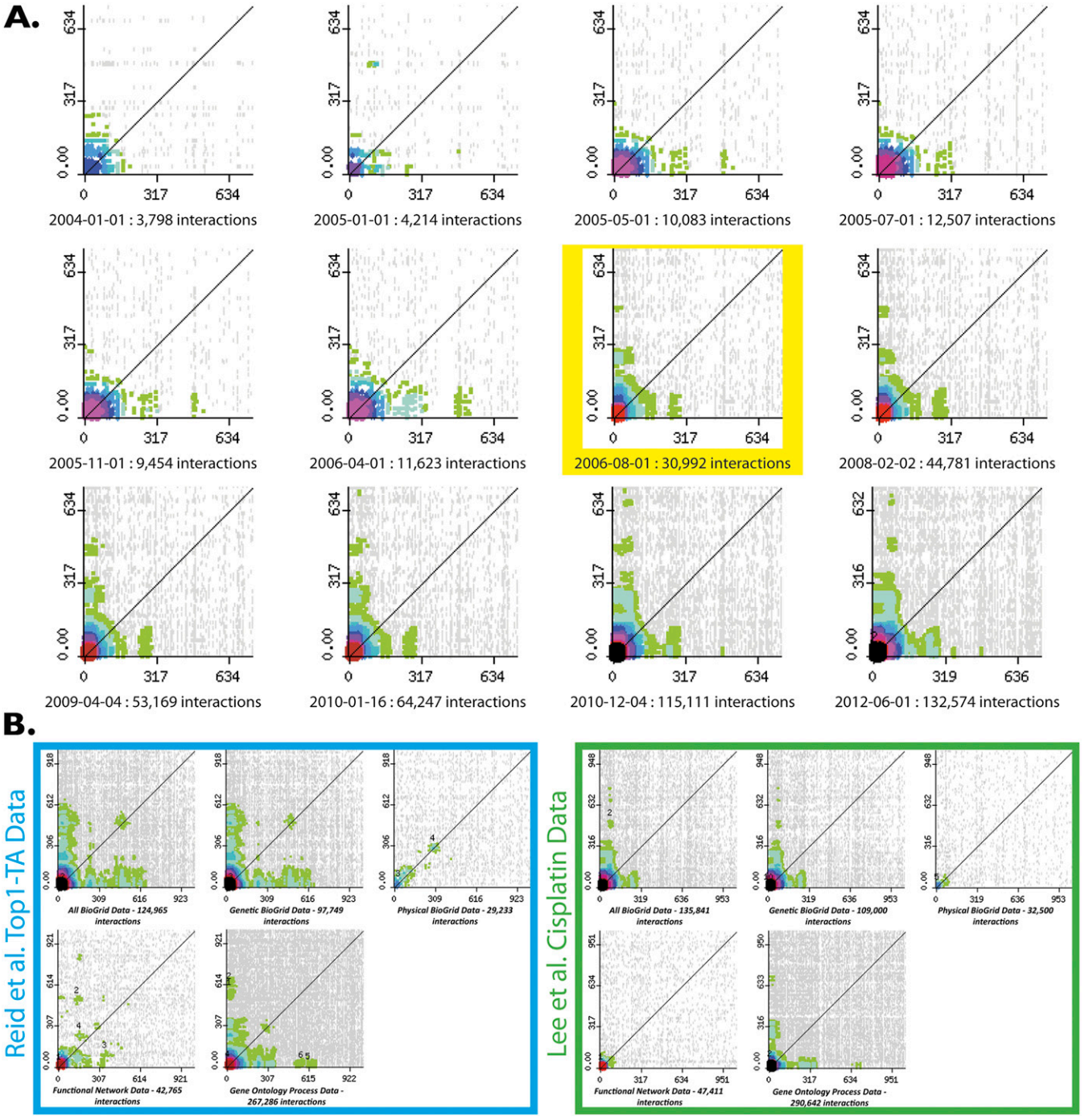

Fig. S6. (*A*) CLIK analysis of cisplatin scrateen results using archived versions of the BioGRID database. The rank-order list of the cisplatin screen was subjected to CLIK analysis using archived versions of the BioGRID database. Graphs should be read by row. The first seven graphs show the first seven sets of *Saccharomyces cerevisiae* data released by the BioGRID consortium. For the next five analyses, every 10th release by BioGRID is shown. After the seventh point (shown by the yellow box), the CLIK group shape remains relatively unchanged. To save space, only the top 700 mutants in the rank order are shown. (*B*) *top1-T*$_{722}$*A* and cisplatin CLIK graphs generated using different sources of interaction data. The interaction databases used were as follows: all data within BioGRID (the default setting when running CLIK), only physical data from BioGRID, only genetic data from BioGRID, interactions annotated in the YeastNet functional interaction database (1), and "interactions" annotated by GO Process terms (2). CLIK graphs truncated to only show the top ~1,000 most sensitive mutants in each screen. The data source and number of interactions in each CLIK graph is indicated below each plot.

1. Lee I, Li Z, Marcotte EM (2007) An improved, bias-reduced probabilistic functional gene network of baker's yeast, Saccharomyces cerevisiae. *PLoS One* 2(10):e988.
2. Ashburner M, et al.; The Gene Ontology Consortium (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25(1):25–29.

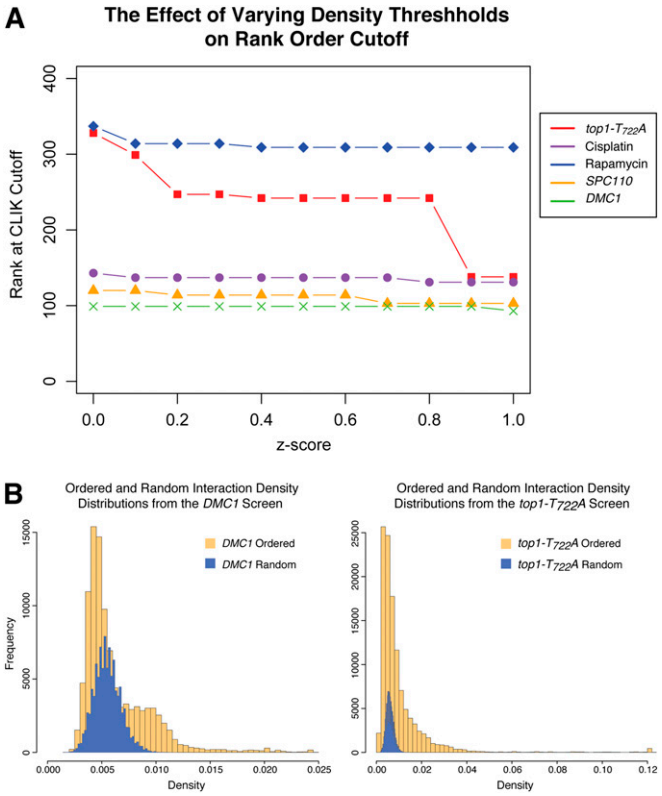

Fig. S7. (*A*) Effect of modifying the number of SDs from the mean on CLIK cutoff selection. For all screens indicated, the number of SDs from the mean in Eq. **S1** was varied from 0.0 to 1.0. (*B*) Overlay of histograms showing the random and ordered density distributions as calculated by the CLIK algorithm.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)