Supporting Information

Decourty *et al.* 10.1073/pnas.0710533105

SI Methods

Yeast Strains and Plasmids. A diploid strain expressing Scd6-eGFP and Dhh1-mRFP fusions was obtained by mating the SCD6-eGFP/HIS3 strain derived from BY4741 (*MAT***a** *leu2Δ0 lys2Δ0 ura3Δ0 met15Δ0*) (1) with the DHH1-mRFP strain (*MATα*) obtained by homologous recombination insertion of the mRFP/KanMX4 cassette at the 3′ end of *DHH1* (BY4742) (2).

The strains used for mRNA decay experiments derived from the RNA polymerase II defective *rpb1-1 ts* mutant strain DLY559 (*MATα ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 rpb1-1*) that allows shutting down global RNA polymerase II transcription with a temperature shift (3, 4). The strains carrying the *SCD6* or *PBP1* deletion in *rpb1-1* mutant strain, were obtained by replacement of *SCD6* or *PBP1* with the KanMX4 cassette whereas *EDC3* was replaced by the prMFα2Nat^R cassette to construct the *edc3*Δ *scd6*Δ and *edc3*Δ *pbp1*Δ double mutants. The plasmid pRP485 (5) was transformed into the mutant strains to express the MFA2pG modified mRNA reporter.

The strains expressing either *PBP1* or *EDC3* under the control of the Tet O2 promoter were obtained by the strategy outlined by (6), using the pCM224 plasmid as template for PCR amplification. The genotypes of the initial strains are *yRP1345:MATα trp1 leu2-3,112 lys2-201 ura3-52, dcp1-2:TRP1 ski8*Δ::*URA3* (7) and *yRP1194:MAT***a** *trp1 leu2-3112 his4-539 lys2-201 ura3-52 ski8*Δ::*URA3* (8).

The pAS2ΔΔ-Scd6 plasmid for the two-hybrid screen was obtained by Gateway cloning in the pAS2ΔΔ destination vector.

Detailed Protocol for Genetic Interactions Mapping (GIM) Screens and Data Analysis. ***Construction of query mutation strains.*** The query nonessential gene deletion strains were generated by homologous recombination with a DNA fragment containing the *MATα* haploid-specific nourseothricin resistance cassette (prMFα2-Nat^R) using the corresponding BY4742 (*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*) derived strain from the yeast knockout mutant collection (9). This fragment was obtained by AscI and EcoRI digestion of the pGID3 plasmid and contains regions of homology with the kanamycin resistance cassette in the *TEF* promoter and terminator regions (see Fig. S1). One microgram of the digestion product was used for transformation of the cells. The transformants were selected on YPD plates containing 0.06 mg/ml nourseothricin (ClonNAT). Selected colonies were next tested for resistance to kanamycin on YPD plates containing 0.2 mg/ml G418. The reference strains (*ymr326c*Δ::*prMFα*2Nat^R *yel068c*::*prMFα2Nat*^{R} and *yfr057w*Δ::*prMFα*2Nat^R) were obtained by the same strategy.

Next, the obtained *MATα* strains were transformed with the pGID1 plasmid bearing hygromycin resistance, used in the next step as a selection markers for the diploid double mutants (hygromycin 0.2 mg/ml in YPD medium) (10).

Double mutants, heterozygous diploids formation. A pool of viable deletion mutants was obtained by growing the entire collection of *MAT***a** cells from the collection to saturation in 96 wells microplates. Aliquots of the pool were mixed with glycerol (25% wt/vol final concentration) and stored at −70°C. Each aliquot of the pool used for a screen contained ≈10^4 cells of every mutant. For every mating with the query mutation strain, the pool of mutants was thawed and incubated for 30 min in GNA medium (5% D-glucose, 3% Difco nutrient broth, 1% Difco yeast extract) at 30°C with agitation and then mixed with cells from 25 ml of a fresh culture of the query strain in fresh, rich GNA medium containing 0.2 mg/ml hygromycin ($OD_{600\ nm}$ around 0.6). For each experiment series, two reference populations were generated and processed in parallel. After a short centrifugation, the mixes of opposite mating type cells were resuspended in 0.5 ml GNA plus hygromycin medium and spread on a 90 mm diameter GNA plate. After 5 h of incubation at 30°C, the cells were recovered from the mating plate and resuspended in 100 ml GNA plus hygromycin plus G418, then left at 30°C with agitation for 24 h.

Sporulation and double mutants haploid selection. Sporulation was performed essentially as described in the Saccharomyces Genome Deletion Project web pages protocol. The selected diploid cells grown in GNA rich medium were washed twice with sporulation medium supplemented with amino acids [1% potassium acetate, 0.005% zinc acetate, uracil (10 mg/l), histidine (10 mg/l), leucine (20 mg/l)] and diluted to a total volume of 100 ml, at final $OD_{600\ nm}$ 0.8. The culture was incubated for 6 days at 25°C and 4 days at 30°C under continuous agitation. Half of the sporulated culture was recovered by centrifugation, resuspended in 250 ml YPD medium. After incubation for 5 h at 30°C with agitation to allow spore germination, G418 and nourseothricin were added for a final concentration of 0.2 mg/ml and 0.06 mg/ml, respectively. To follow the population of selected double mutant haploid cells over a number of 18 generations we used initially successive dilution of cultures when the $OD_{600\ nm}$ reached values higher than 1.5. Better reproducibility was obtained when the cultures were grown in an automatic turbidostat (machine described in ref. 11). Estimates of growth speed were based on absorbance readings or on cumulated volume of diluting medium when the turbidostat was used. In many experiments, counts of viable double mutant haploids were done by spreading different volumes of culture on YPD plates containing G418 and nourseothricin. The equivalent of 5 $OD_{600\ nm}$ units of cells was recovered by centrifugation after completion of an estimated number of 18 generations and frozen at −70°C.

Tags amplification and microarray hybridization. Total DNA was extracted from the cells using a phenol-chloroform, glass beads protocol. For each sample, cells were suspended in 0.2 ml extraction buffer (2% Triton X-100, 1% SDS, 0.1 M NaCl, 10 mM Tris·HCl pH8, 1 mM EDTA) to which ≈0.2 ml acid-washed glass beads (0.4–0.6 mm diameter, Sigma) and 0.2 ml mix phenol:chloroform:isoamyl alcohol (25:24:1) pH 8 were added. After thorough vortexing (7 min), 0.2 ml TE buffer was added and the samples were vortexed briefly. After 10 min centrifugation at maximum speed at room temperature, the supernatant was mixed with 1 ml ethanol. Maximum speed centrifugation at 4°C was followed by a brief wash of the pellet with 70% ethanol. Finally, the pellet was suspended in 30 μl TE with added RNase A and kept frozen at −20°C.

Tag amplification was performed in two successive steps, the first being specific for the tags flanking the kanamycin resistance cassette (see Fig. 1). For each sample two different series of PCRs were done, one for the upstream tags and another one for the downstream tags. Sequences of the primers used were: KU: AAG AAG AAC CTC AGT GGC, U1: GAT GTC CAC GAG GTC TCT, U2: CGT ACG CTG CAG GTC GAC, KD: GGA TCT TGC CAT CCT ATG, D1: CGG TGT CGG TCT CGT AG, D2: ATC GAT GAA TTC GAG CTC G.

The first PCRs were done with primers U1+KU and KD+D2, using 25 cycles of amplification, with an annealing temperature of 50°C and extension time of 30 s. The template was 1 μl of the genomic DNA preparation from the previous step (in standard

50-μl PCRs). Primary PCR products were purified using the QIAquick kit (Qiagen).

The query and control PCR products were differentially labeled by performing secondary PCRs using Cy3 and Cy5 5′-end-labeled versions of U2 and D2 primers respectively. The volume for the secondary PCR for one tag was 100 μl, using 2 μl of the purified primary PCR product, 55°C as annealing temperature and 15 s of extension time (15 cycles). Antisense unlabeled primers (2 μM final concentration, mix of U1, D1, antisense U2 and antisense D2) were used to suppress nonspecific hybridization. To the final mix of 400 μl of 4 PCR products (UPTAG and DOWNTAG PCR from the query deletion screen and from the reference screen), sodium acetate, pH 5.2 was added to a final concentration of 0.3M and the DNA was precipitated by addition of 1.1 ml 95% ethanol in presence of 10 μg linear polyacrylamide. After 1 h precipitation at −20°C, the pellet was recovered by centrifugation, washed with 70% ethanol, suspended in 5 μl water and mixed with 100 μl DIG Easy Hyb hybridization buffer (Roche).

The samples were hybridized to glass slides oligonucleotide microarrays from SLRI/Mount Sinai Hospital, Toronto (SLRI_Yeast_Barcode_13k, EBI ArrayExpress accession number A-MEXP-714). Samples in hybridization buffer were denatured for 2 min. at 95°C and cooled on ice for 5 min. After an incubation for several min. at 55°C, samples were applied to a glass slide microarray and left at 25°C overnight. Hybridizations with query DNA Cy3-end-labeled version and the two different reference DNA Cy5- end-labeled (*ymr326c*Δ::*prMFα*2Nat^R and *yfr057w*Δ::*prMFα*2Nat^R or *yel068c*::*prMFα2Nat*^{R}) were run in parallel. Routinely, the experiments were done with a set of 6 query mutation strains and two reference strains.

The slides were washed in 6× SSPE, 0.05% Triton X-100 for 5 min, and in successive baths of 2× SSPE, 0.05% Triton, 0.2× SSPE, 0.05% Triton and 0.2× SSPE, for 3 min. each and dried by a short centrifuge spin. Alternatively, the slides were washed in a Tecan HS 4800 hybridization station using the same buffers as above. Scans of the microarrays were done on a GenePix 4000B scanner using the GenePix Pro software. The images were analyzed and obvious artifacts were eliminated manually. The obtained values were exported in text file format and stored in a PostgreSQL database. The raw results were normalized by the "loess" algorithm for each microarray hybridization using R and the BioConductor package (12). The signal/noise ratio and the diameter of the quantified spots were used to filter out hybridization artifacts. For each deleted ORF, the median of the values for the different spots corresponding to the corresponding UPTAG and DOWNTAG was calculated. Aggregated results were exported to comma delimited text files and further processed using programs developed using the Python and Ocaml languages.

Exclusion peaks correction. The relative measured fitness of the double mutants from each screen, estimated from the values of the intensity of the signal of the hybridization on microarrays for the query double mutant population compared with the reference double mutant, log2(Q/R), were plotted against their relative physical position in the genome sequence (see Fig. S1*D* for an example). Two peaks were observed; one peak was centered on the position of the query gene deletion while a second one was centered on the position of the reference deletion. To obtain haploid double mutants through a meiotic division, recombination by crossing-over for loci that are on the same chromosome is required. As a result, the proportion of double mutants for deletions of ORFs that are close on the chromosome will be lower than that obtained for ORFs that are far apart or on different chromosomes. This lower proportion of recombinants will lead to a lower signal level on the microarrays when screens with two different query deletions are compared (deletion of interest compared with reference deletion). The normalized ratio query/reference thus provides an estimate of the recombination frequency, if no specific genetic interactions interfere.

An overview of the average signal ratio (query screen/reference screen), log2 transformed, obtained for deletions at different distances from the query deletion position is shown in Fig. S2. A log2(Q/R) of 0 indicates no genetic linkage between the two antibiotic resistance markers that replaced the deleted ORFs. A value of −1 for log2(Q/R), ratio of 0.5, indicates that only 50% of the signal for a given double mutant was recovered when compared with the double mutant cells obtained with the reference gene. To obtain a linear relationship between the recombination frequencies and physical distance that would allow us to correct or eliminate the values, we used the Haldane mapping function

$$\mathbf{r} = 1/2(1 - e^{-2\mathbf{m}}); \quad \mathbf{m} = 1/2(-\ln(1 - 2\mathbf{r})),$$

where **r** is the recombination frequency and **m** is the genetic distance. The expected recombination frequency for unlinked markers is 0.5. An equivalent of recombination frequency **r**′ can be obtained from the experimental values of ratios of signal in the query mutant screen versus the reference screen

$$\mathbf{r}' = 1/2^{\text{ratio}},$$

where the sign of the log2(ratio) was reversed, if necessary, to give values that are greater than 1 for the peak that was analyzed. Values of ratio equal or inferior to 1 indicate no linkage and could not be used in calculating the corrections.

The Haldane transformation of **r**′ should establish a linear correlation between the result, equivalent of genetic distance, and the physical distance between the two markers on the chromosome. We considered the distances between the ORF centers coordinates (as annotated in the Saccharomyces Genome Database, www.yeastgenome.org) as approximations of distances between the two antibiotic resistance markers. For the averages of 119 hybridizations, on 41 different loci, there was a good linear correlation between the genetic distance **m** and physical distance (Fig. S2 *Inset*).

To correct individual exclusion peak values, a smoothing step by the moving average method over 7 values of recombination frequency equivalent (**r**′) was first applied. Next, the Haldane mapping function was applied to the smoothed curves and a line was fitted through the data using only the interval of 0.1–0.75 map units. This way, an estimate of the **r**′ values as a function of physical distance, was obtained. The linear regression parameters were used to adjust every individual **r**′ value by the difference between 0.5 (no linkage) and the calculated value from the distance between the markers:

$$\mathbf{r}'_{\text{corr}} = \mathbf{r}' + 0.5 - \mathbf{r}'_{\text{calc}}.$$

To avoid over-corrections, a maximal value of 0.79 for $\mathbf{r}_{\text{corr}}$ was used. Finally, the results were retransformed to ratios, and then to log2(Q/R).

A region of 15 kb on either side of the position of the deleted query gene and reference deletion were excluded from the analysis because the values obtained at such distances were not reliable to assess recombination rates. If no linear regression could be computed, values from an arbitrary region of 200 kb on the side on which the correction failed were excluded from further analyses. Images of the exclusion peaks before and after correction were verified for each correction and the values that were considered biased were manually eliminated.

The corrected values were used to test the reciprocity of the observed genetic interactions. When a prey was found in a given screen and was also used as query in another screen, we plotted

the log2(Q/R) observed in the two screens for the same pair of mutants (Fig. S5*B*).

Filtering results to account for specificity toward a pathway. Once the peaks of values corresponding to genetically linked deletions were corrected, the log2(Q/R) values were merged in a single table and two successive reproducibility filters were applied. First, for most experiments, the hybridization results with two different reference deletions were compared for each screen, when available. If only one value for a given ORF was available from two hybridizations, the value was dropped. The values were sorted and the corresponding ORFs ranks were normalized to range between 1 and 4,700. We next calculated the difference in rank for every measured ORF in the two hybridizations. For experiments where only one reference screen was available all of the ORF values were kept for further filtering.

In search for a method to filter out the data that were not reproducible, we visually analyzed several data sets to establish rules that would allow automatic removal of results that we considered irreproducible, especially for cases of high absolute values. Such a function was based on the ranks of the values obtained in two independent screens. For each ORF, the rank differences between two hybridizations were raised to power 7, compared with a threshold value and only differences that were inferior to the threshold (2,300 raised to power 7) were kept for further processing. This way, the values of low amplitude where not affected by filtering and the corresponding results, mainly indicating a lack of genetic interaction, which by itself was an important result, were preserved in the reported data.

In a second filtering step, we used the same strategy to compare experiments performed independently with the same bait, using the same difference in rank strategy but with more stringency, the arbitrary threshold being set to $2{,}100^7$. Only double mutant relative fitness values that were present after applying the two filters in at least half of the 73 independent screens and showing a difference of at least 2 in at least one experimental condition were used. Hierarchical clustering of the normalized, average log2 transformed 1,095 values was performed with Cluster 3 software (13) using uncentered Pearson correlation and hierarchical clustering of both rows and columns.

Results, adjustments, and comparison with protein–protein or GO annotations. Deletions showing similar fitness changes when combined with the 41 tested deletions corresponded to genes that either code for directly interacting proteins or are involved in a common cellular pathway. The similarity between the fitness effects on different double mutants was measured using the Pearson correlation coefficient. The correlation coefficients were calculated for every possible combination of the 1,095 selected deletions. Random shuffling of the values from each row generated a control table. For the obtained correlation values, the R statistical software (14) was used to obtain the frequency distribution (histogram) of Pearson coefficients (Fig. 4*B*).

Two comparisons were done to investigate the ability of the correlations in predicting protein–protein or functional interactions. First, we generated the list of overlapping deletions present in the 1,095 selected ones. If two predicted deleted ORFs were at a distance of less than 10 nucleotides, the two deletions were presumed to have a similar effect. A total of 51 pairs of such overlaps were present and the distribution of the corresponding Pearson correlation coefficients was compared with the distribution of all of the pairs. Second, we extracted high confidence protein interaction data from the metastudy presented in ref. 15 and only considered pairs with a specificity score (PE) superior to 3.19. From these, only the 218 pairs that contained ORFs from the 1,095 ORFs set were selected and the corresponding distribution of correlation coefficients was compared with the distribution of the values for all of the pairs.

The "gold standard" of Gene Ontology (GO) annotation defined in ref. 16 on the basis of the usefulness of a chosen subset of 200 GO terms was used to estimate the precision and coverage of the data. For example, we took all of the ORFs pairs showing profiles with Pearson correlation coefficients larger than 0.3 and selected the 46,170 "gold standard" corresponding pairs (5,182 positive and 40,988 negative). The "gold standard" pairs accounted for 680 ORFs being in "true" pairs and 780 ORFs being in "false" pairs. We counted the number of identified true positive (TP) pairs and false positive (FP) pairs at different levels of Pearson correlation coefficient and computed a "specificity" score as TP/(TP + FP). The number of different ORFs that were found in TP pairs was used as a measure of the coverage. This approach allowed us to improve the specificity and coverage of the results by fine tuning the score for the strength of correlation while taking into account the distribution of the correlation coefficient values for every ORF when paired with all of the others. A symmetrical, short-tailed distribution was typical for most of the ORFs, while long-tailed, asymmetrical distributions were characteristic for ORFs that showed specific and informative pairing with other ORFs. We used kurtosis as a global measure of the flatness of the distribution and corrected the Pearson correlation coefficient with an empirical equation:

$$\text{pearson}_{\text{corrected}} = \text{pearson} \times (\text{abs}(\text{pearson}))^{\mathbf{f}},$$

where

$$\mathbf{f} = \mathbf{n}/\text{kurtosis}^{\mathbf{x}},$$

with **n** and **x** arbitrarily chosen to be 15 and 4, respectively. The net result of this transformation was an increase in the values of the correlation coefficients values for ORFs that showed specific associations while decreasing the correlation coefficients for ORFs showing many nonspecific associations (as assessed by comparison of the results with the "gold standard" of known functional associations). The transformed Pearson correlation coefficients were used to estimate the precision and coverage of the results. Three classes of association were defined on the basis of the GO "gold standard": strong (256 pairs), medium (573 pairs) and weak (1,190 pairs) association with predicted precision levels of over 0.7, 0.6 to 0.7 and 0.5 to 0.6, respectively.

Specificity scores for combined fitness defects (SGD score calculations). While similar genetic interaction profiles (GIP) were useful in uncovering novel functional interactions, the original purpose of the GIM screens was the identification of synthetic growth defects (SGD). Because some of the deletions were identified in a large number of screens, knowledge about such broad fitness defects was of little use for describing specific functional associations. To minimize the influence of such results on the final estimates of genetic interaction we set up an optimization strategy that took into account four parameters. First, we took the log2(Q/R) measured when combining a query deletion with a hit deletion as the main factor (LQR). Second, for each value in a genetic interaction profile we derived a "median log2(Q/R)," calculated as the median of the log2(Q/R) values that had the same negative sign as LQR. The values corresponding to query gene screen of the same "group" were not taken into account (see Table S6 for group association of the query genes). This "median" parameter would be strong and negative if the gene of interest showed a synthetic growth defect with most of the query genes deletions. Third, we used the logarithm of the kurtosis of the distribution of log2(Q/R) values from a hit deletion, where only values from an interaction profile that were weaker than the LQR were considered. Fourth, a rank score that decreases very rapidly with increasing rank was used to increase the importance of the deletions that had the highest effect in a given screen, even if this effect was marginal when compared with other screens.

The rank score was set at 0 if the rank was superior to 20 and as $1/2^{rank}$, if the rank was between 1 and 20.

We manually selected 151 pairs of gene deletions associated by our screens, considered specific, and used them as a set of "true" pairs. The number of considered pairs was limited to the best synthetic growth defect 2,500 values for each screen, giving a total of 35,796 measurements. A logistic regression model was fitted to these data sets using the data mining software Weka (17) to find out optimal weights for the four parameters in defining "specific" genetic interactions (the coefficients obtained were −1.4541, 0.955, 1.7773 and 0.6431). A score (SGD score) was calculated using the sum of the products of coefficients and parameters and the list of results was compared with the gold standard. Because only a low number of pairs were available from these results in the GO gold standard the precision and coverage estimates were only rough approximations (Fig. 4*C*).

Interaction network combining GIP and SGD. It was obvious from the results of the above described data manipulations that we would gain more information from our screens if we could combine the results of direct SGD and GIP. To build a reliable interaction network that could be explored with software tools like Cytoscape (18) we combined strong GIP and strong SGD in a single list of high confidence interactions. Next, this network was extended using the medium confidence results, adding only those interactions for which one of the genes was already present in the high confidence network. Last, for subnetworks of interest, we added the additional information from low confidence pairs, only if both components of a pair were part of the subnetwork. To find GO terms enriched in subnetworks we used GO Term Finder on the SGD web site (www.yeastgenome.org) and annotated groups of genes as represented in Fig. S7.

Other Methods. TAP purifications were performed from 4 liters of yeast culture as described in (19) at lower stringency, using buffers containing 0.1M NaCl. The purified complexes were separated on a 5–20% polyacrylamide gradient-SDS gel and protein identifications were done using MALDI-TOF mass spectrometry. The yeast two hybrid screen was performed using a cell to cell mating strategy as previously described (20). Fluorescence microscopy. Diploid cells expressing both Scd6-eGFP and Dhh1-mRFP fusion proteins were grown in synthetic media containing glucose up to $OD_{600\ nm}$ 0.5 then recovered by centrifugation, washed and resuspended in synthetic media lacking glucose for 15 min. *In vivo* mRNA assay for MFA2pG degradation experiments were performed as described in (5). The estimation of *TetO2 pbp1* or *TetO2 edc3* mutants growth was done by serial dilutions of initially equal numbers of cells, on rich medium plates with or without 4 μg/ml doxycycline, followed by an incubation at 25°C or 30°C for 24 h.

1. Ghaemmaghami S, *et al.* (2003) Global analysis of protein expression in yeast. *Nature* 425:737–741.
2. Longtine MS, *et al.* (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* 14:953–961.
3. Wyers F, *et al.* (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* 121:725–737.
4. Nonet, M., Scafe, C., Sexton, J., Young, R (1987) Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Mol Cell Biol* 7:1602–1611.
5. Decker CJ, Parker R (1993) A turnover pathway for both stable and unstable mRNAs in yeast: Evidence for a requirement for deadenylation. *Genes Dev* 7:1632–1643.
6. Belli G, *et al.* (1998) An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Res* 26:942–947.
7. Dunckley T, Parker R (1999) The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. *EMBO J* 18:5411–5422.
8. Anderson JS, Parker RP (1998) The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. *EMBO J* 17:1497–1506.
9. Winzeler EA, *et al.* (1999) Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science* 285:901–906.
10. Berger AB, *et al.* (2007) Hmo1 is required for TOR-dependent regulation of ribosomal protein gene transcription. *Mol Cell Biol* 27:8015–8026.
11. Frachon E, Bondet V, Munier-Lehmann H, Bellalou J (2006) Multiple microfermentor battery: A versatile tool for use with automated parallel cultures of microorganisms producing recombinant proteins and for optimization of cultivation protocols. *Appl Environ Microbiol* 72:5225–5231.
12. Gentleman RC, *et al.* (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol* 5:R80.
13. de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. *Bioinformatics* 20:1453–1454.
14. R Development Core Team (2006) R: A Language and Environment for Statistical Computing. Vienna, Austria (www.r-project.org).
15. Collins SR, *et al.* (2007) Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. *Mol Cell Proteomics* 6:439–450.
16. Boyle EI, *et al.* (2004) GO:TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20:3710–3715.
17. Witten IH, Eibe, F (2005) *Data Mining: Practical machine Learning Tools and Techniques* (Kaufmann, San Francisco).
18. Shannon P, *et al.* (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504.
19. Rigaut G, *et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17:1030–1032.
20. Fromont-Racine M, Rain JC, Legrain P (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* 16:277–282.

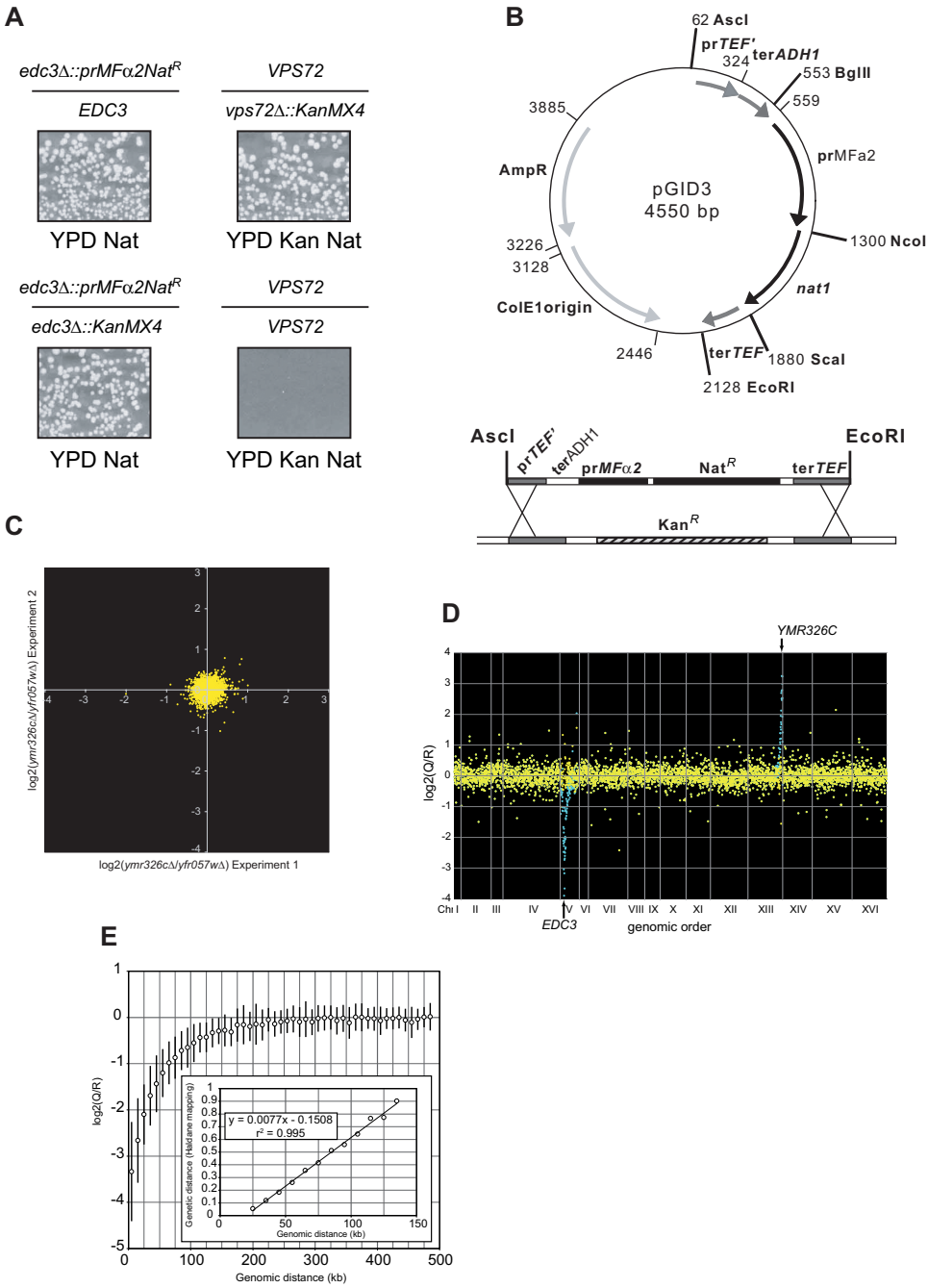

Fig. S1. Tools for the GIM screens. (*A*) Efficient selection of haploid double-mutants using the haploid-specific nourseothricin resistance cassette. Sporulation was induced for diploid heterozygous deletion cells where one copy of *EDC3* was replaced by the prMFα2-Nat^R cassette and a copy of *VPS72* was replaced by the Kan^R cassette. Cells were spread on YPD plates containing antibiotics. Both single selection for Nat^R and double selection for Nat^R/Kan^R allowed growth of, respectively, haploid (*MAT*α) cells and haploid (*MAT*α) double mutant cells (*Top*). A similar experiment was done with both markers at the same locus, *EDC3*. Formation of haploid double-mutants is impeded by the presence of both markers at the same locus, thus no growth was observed on medium containing both selection antibiotics (*Lower*). This demonstrates the efficiency of the selection for the haploid cells and the counter selection of diploids. (*B*) The downstream half of the pr*TEF* promoter of the pAG25 plasmid [Goldstein AL, McCusker JH (1999) *Yeast* 15:1541–1553] (harboring the nourseothricin resistance gene *nat1*) was replaced by the *ADH1* terminator followed by the haploid specific promoter from the *MFα2*/*YGL089C* gene to generate plasmid pGID3 (*Upper*). The reconstituted sequence of the plasmid can be obtained on request. Digestion of this plasmid with AscI and EcoRI generates a DNA fragment that is used to replace, by homologous recombination, the kanamycin resistance cassette by the haploid-specific nourseothricin resistance cassette (*Lower*). Homologous recombination occurs at the level of the *TEF* promoter and terminator sequences remaining in common between both cassettes. (*C*) Control of the neutrality of reference strains by comparison of two independent screens with *ymr326c*Δ as a query mutation against *yfr057w*Δ as a reference strain. Same as in Fig. 2*A*. (*D*) The distribution, ranked by the physical position of deletions along the chromosomes, of the log2 of the ratios between the signal intensities [log2(Q/R): median of the log2(Q/R) values for the filtered duplicated UPTAG and DOWNTAG] for each mutant in the *edc3*Δ::*prMFα2Nat*^{R} query population vs. the reference population (*ymr326*Δ::*prMFα2Nat*^{R}). Each dot value is proportional to the relative fitness of the double mutant strain in the query vs. the reference population. Vertical bars mark chromosome boundaries. The two arrows point to the genomic locations of *EDC3* and *YMR326C*. Two peaks in opposite directions were immediately evident (blue dots). These peaks were centered on the genomic position of the query and reference genes, spanned 200 kb on average and resulted from the lower occurrence of recombination between two genetic loci if they are on the same chromosome. The observed peaks thus directly reflected the genetic distance between the Kan^R and Nat^R markers {a similar phenomenon is also observed, for the same reason, in the SGA or E-MAP analyses [Collins SR, Schuldiner M, Krogan NJ, Weissman JS (2006) *Genome Biol* 7:R63]}. Aside from these genetic linkage peaks, the negative and positive values corresponded to mutations showing either a synthetic growth defect (SGD) (aggravating interaction) or an epistatic effect (alleviating interaction) with the query mutation. We used two strategies to minimize the number of deletions that had to be excluded from the analyses as a result of their proximity to the query or reference deletion loci. First, we corrected the values in function of the estimated recombination frequencies calculated with the Haldane mapping function (see *E*). Second, two reference deletions were used in independent hybridizations for most experiments. In addition, for most of the query mutations, the experiments were performed twice independently. (*E*) Relation between log2(Q/R) values and genetic distance measured between 41 different query loci and their genetically linked loci. The log2(Q/R) from the query gene screens and reference screens were plotted against the mean genomic distances (kb) between the Kan and Nat markers. Results were obtained from 119 different hybridizations. Because the ratio between the signals in query and control populations are direct indicators of the frequency of crossings-over between two given markers, we used the Haldane mapping function to estimate the genetic distance relation to physical distance on the chromosome. The equation was individually fitted by linear regression on smoothed curves from individual hybridization and used for correction to eliminate this bias.

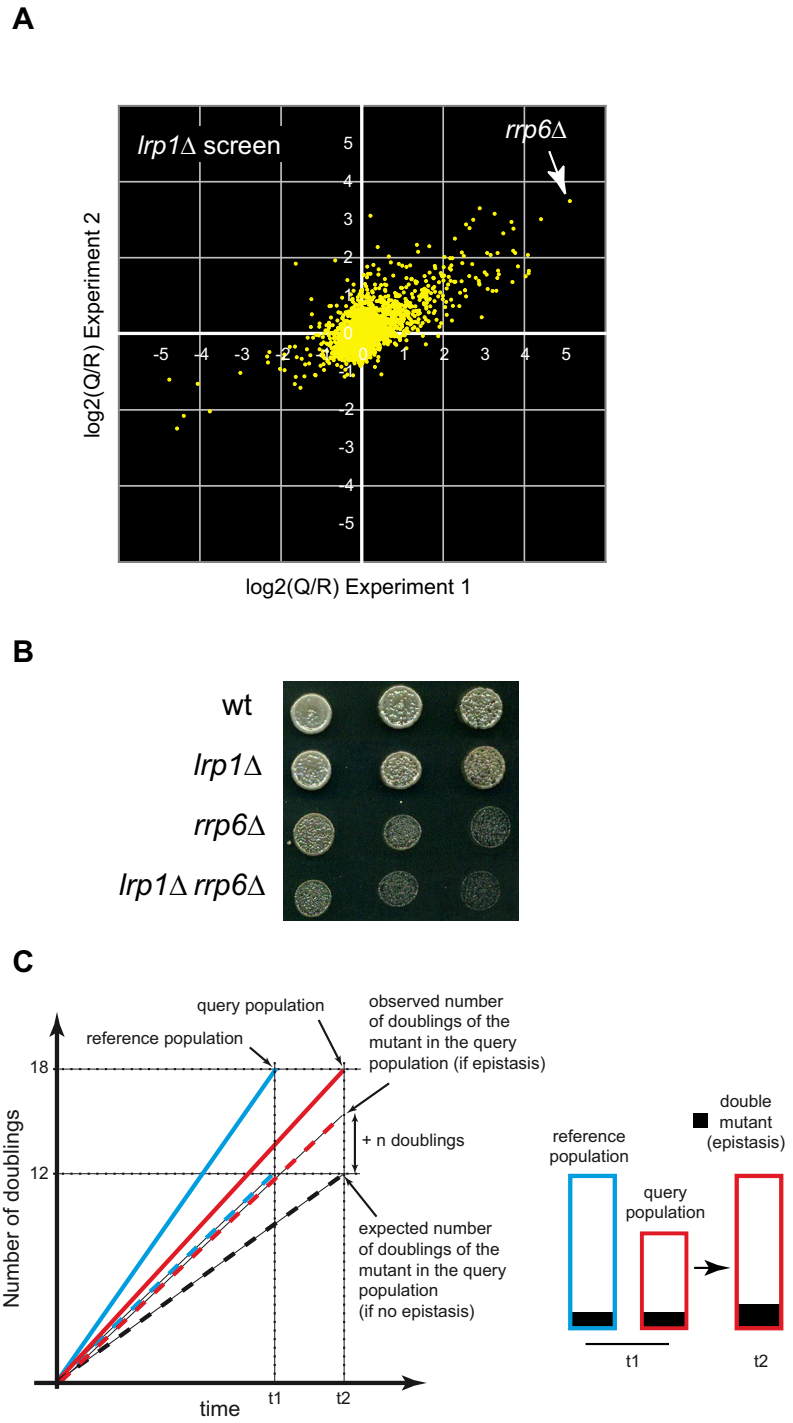

Fig. S2. An example of observed epistasis as a source of positive log2(Q/R) values in the GIM screens (*A*) The results of two independent screens performed with a strain where a component of the nuclear exosome, *LRP1*, was deleted. The highest level of log2(Q/R) was obtained for the double deletion *lrp1*Δ/*rrp6*Δ (indicated by an arrow) (*B*) Serial dilution growth assay on YPD plate to compare wild-type, *lrp1*Δ, *rrp6*Δ and the double mutant growth showing that the double mutant does not grow better than any of the single mutants. Hence, the positive value observed in the screen does not result from a compensatory effect of the single mutants. (*C*) Schematics to explain the occurrence of positive values of log2(Q/R) in the GIM screens with mutants exhibiting a slow growth phenotype and exhibiting epistatic effects. The *Left* shows the doubling rates of a reference population (blue, continuous line) or a query population (red, continuous line) with a slow growth phenotype (*lrp1*Δ in our example). The expected growth rate of a slow growing mutant (*rrp6*Δ in our example) in the *lrp1*Δ query population if the *rrp6*Δ and *lrp1*Δ growth defects were simply additive (no epistasis) is shown as a dashed black line. In fact, *lrp1*Δ being epistatic over *rrp6*Δ, the growth rate of the *rrp6*Δ mutant is similar in the reference population (dashed blue line) and the query population (dashed red line). Because the query population is grown for a longer time (t2) than the reference population (t1), the abundance of *rrp6*Δ relative to the overall population is larger in the query population than in the reference population (see also *Right*).

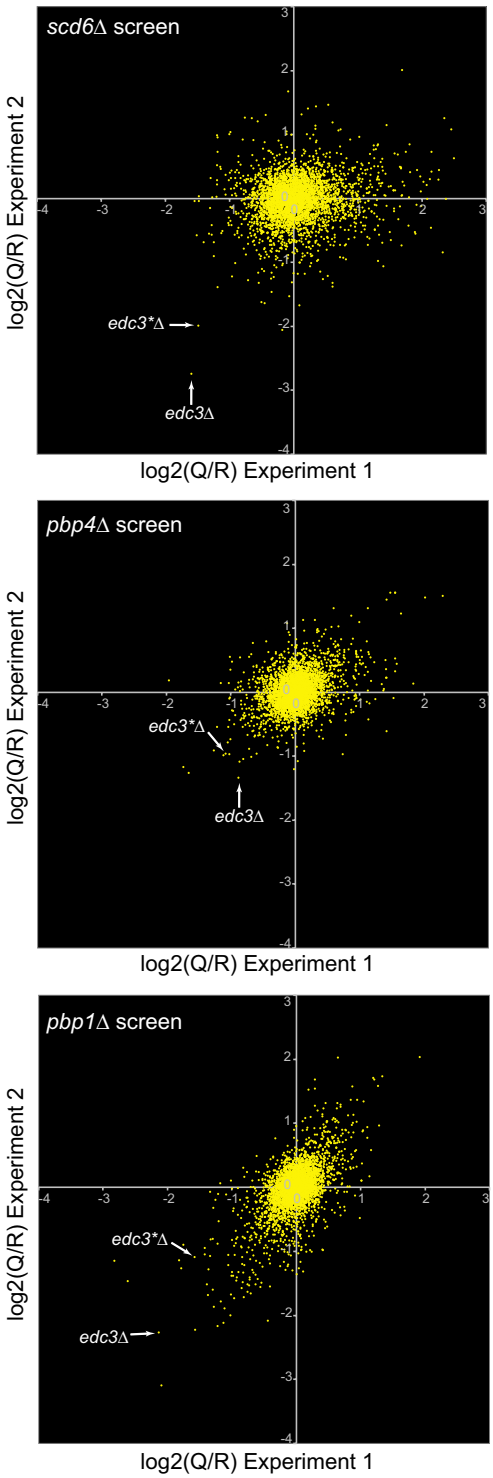

Fig. S3. Reciprocal screens result in reciprocal genetic interactions. Two independent screens with either *scd6*Δ (*Top*), or *pbp4*Δ (*Middle*), or *pbp1*Δ (*Bottom*) as a query mutation are compared. Each dot represents the mean of the ratios obtained with two different reference populations in each experiment. Arrows point to the values obtained for double mutants between *edc3*Δ (or *edc3*Δ*; see Fig. 2*C*) and the query mutation in each screen.

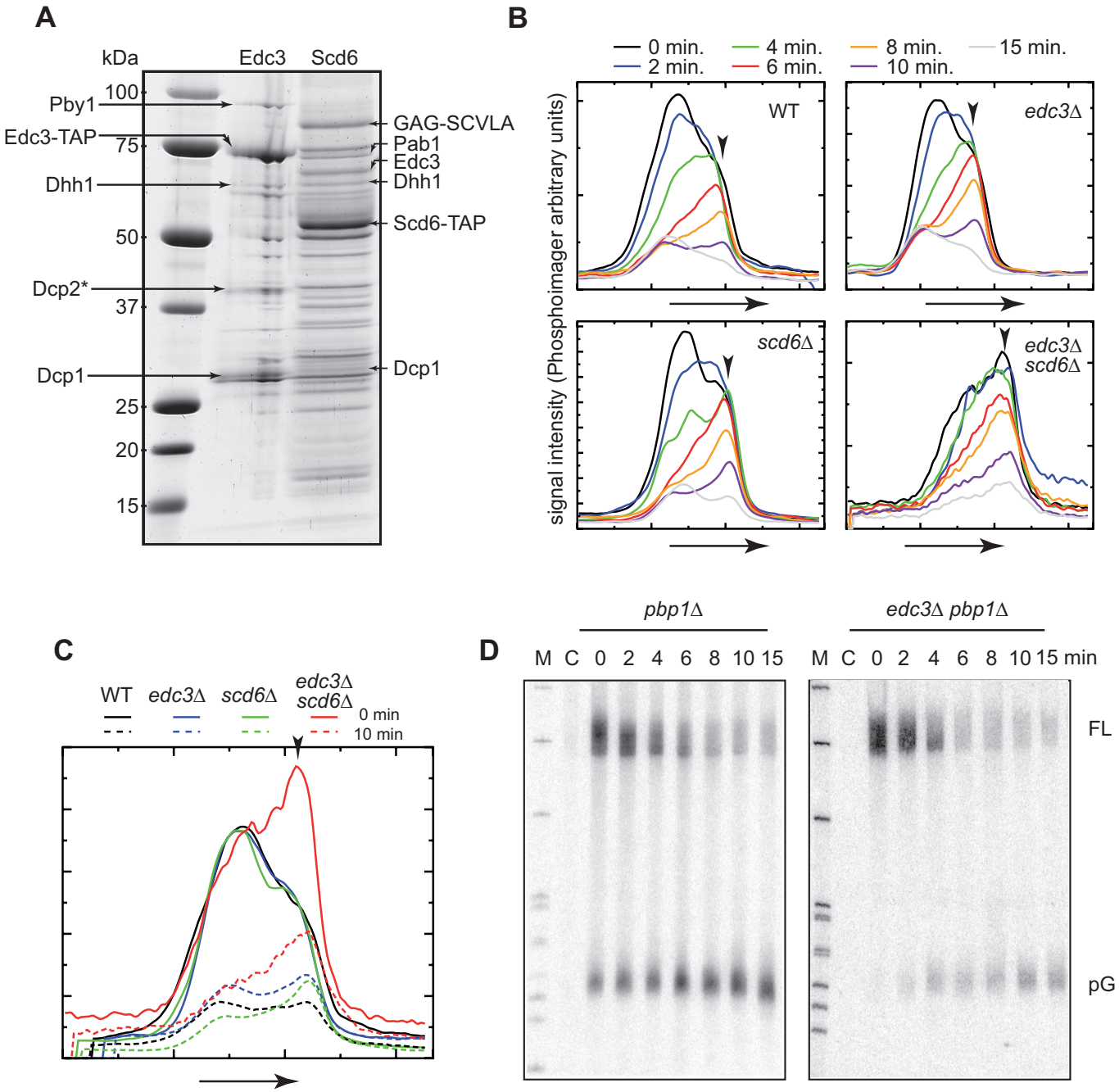

Fig. S4. Supplementary data for functional validations of Scd6 and Pbp1. (*A*) Complexes associated with Edc3-TAP and Scd6-TAP were isolated by tandem affinity purification. After electrophoresis on a polyacrylamide gradient gel and Coomassie blue staining, the proteins present in the visible bands were identified by mass spectrometry (see Table S1, parts *a* and *b*). Arrows indicate proteins that we estimated specific to the purifications. The other non-annotated bands correspond to common contaminants (ribosomal proteins and translation factors). The asterisk indicates a protein fragment. (*B*) Profiles of MFA2pG mRNA deadenylation corresponding to the gels shown in Fig. 3*B*. Phosphoimager scans were quantified using ImageQuant and profiles were generated for each area corresponding to the full-length mRNA in each lane. The vertical arrowheads point to the approximate position of the deadenylated form of the mRNA. The direction of gel migration is indicated by horizontal arrows (*C*) Comparison of the profiles for time 0 and time 10 min. for the wild-type, *edc3*Δ, *scd6*Δ and *edc3*Δ *scd6*Δ strains. Signal intensity was normalized for each strain to obtain similar values for the polyadenylated mRNA fraction. (*D*) MFA2pG mRNA degradation time courses in *pbp1*Δ strains. The experiments were performed in the conditions described in Fig. 3*B*.

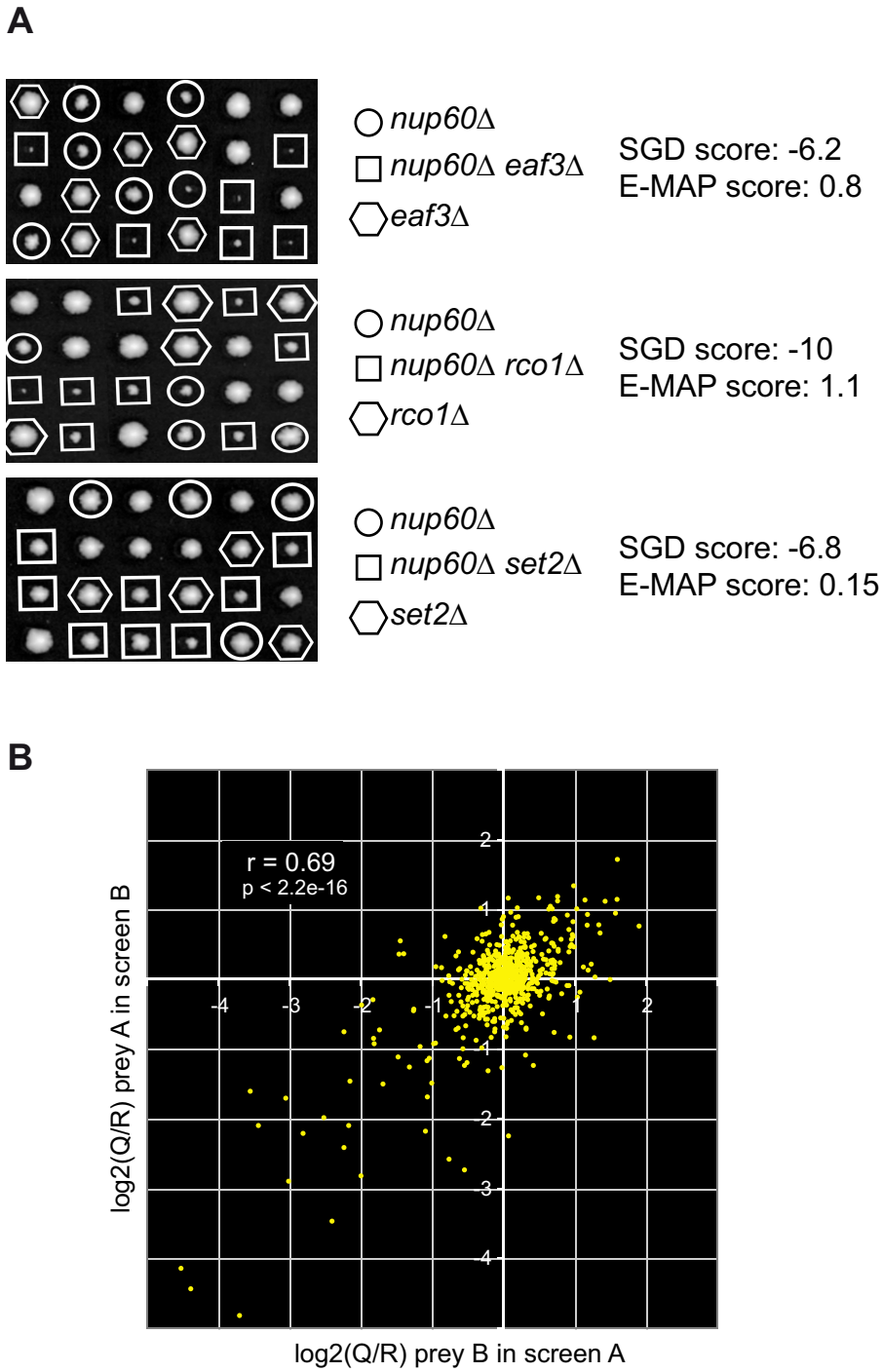

Fig. S5. (*A*) Tetrad analyses of synthetic growth defect mutants identified in GIM and not in the E-MAP methods. For each mutant combinations, six tetrads were analyzed on reach medium. Single and double mutants are as indicated by circles, squares and hexagons. For each double mutants, the SGD scores from the GIM screens and the E-MAP scores are noted on the right. (*B*) The reciprocity of the measured genetic interactions was evaluated by plotting the log2(Q/R) values obtained for prey B in a screen with mutant A against the values obtained for prey A in a screen with mutant B. The corresponding values are provided in Dataset S1.

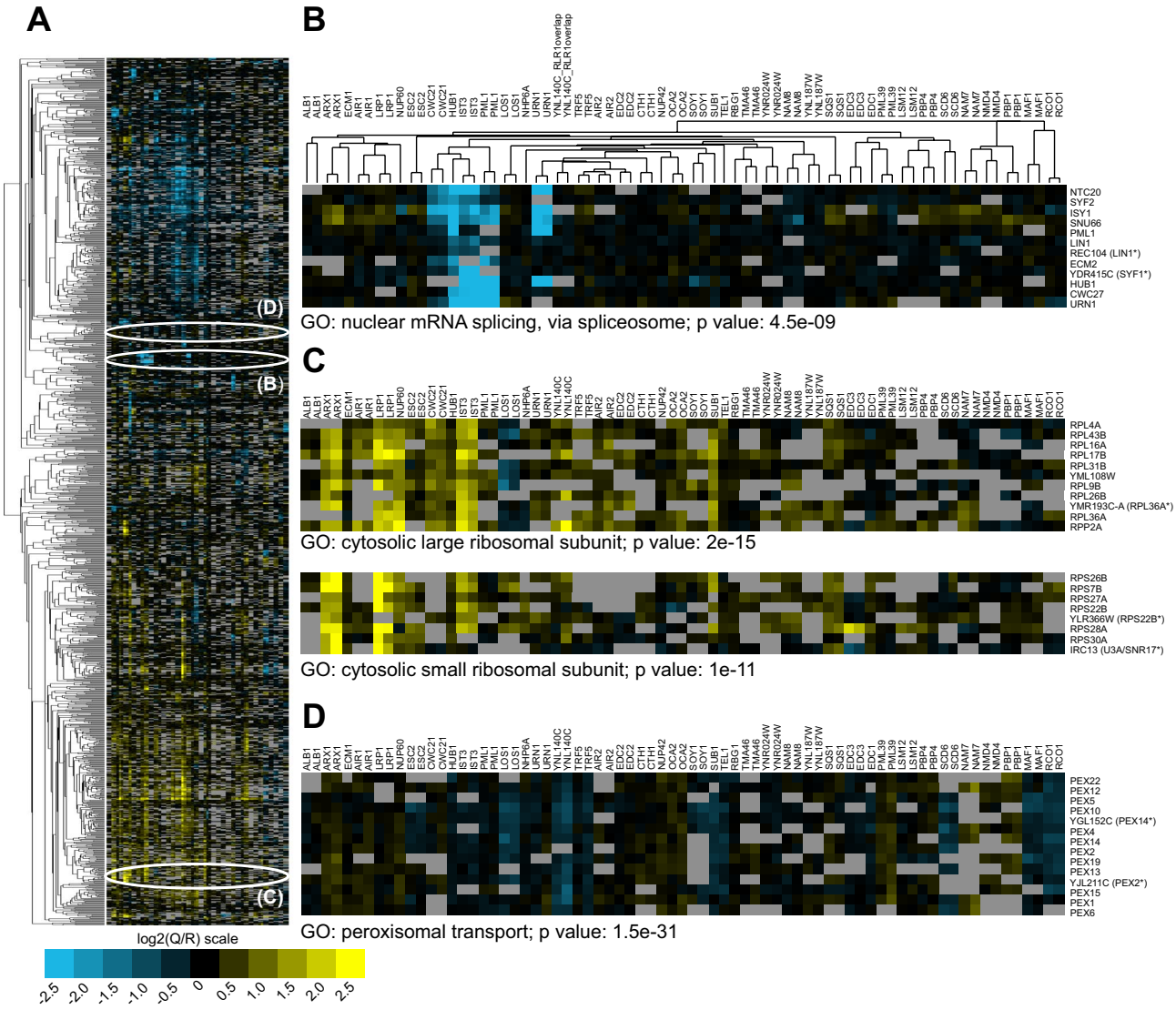

Fig. S6. Global overview of GIM results. A hierarchical clustering of the genetic interaction profiles for the filtered results of 73 independent genetic screens with 41 different query deletions was performed with Cluster 3.0 [de Hoon MJ, Imoto S, Nolan, J, Miyano S (2004) *Bioinformatics* 20:1453–1454] and displayed in Java TreeView [Saldanha AJ (2004) *Bioinformatics* 20:3246–3248]. A detailed view of the whole cluster can be found at www.pasteur.fr/recherche/unites/Gim/SI_decourty.html. Only the 1,095 deletion strains of the collection exhibiting an absolute log2(Q/R) value above one in at least one experiment are represented. The clustering was performed both on target mutations (rows) and query mutations (columns). Pairs of independent screens with identical query genes clustered together, highlighting the reproducibility of the method. (*A*) Overview of the clustering of 1,095 selected hit gene results. Ellipses indicate example cluster regions, enlarged in the other panels of the figure. (*B*) A cluster of splicing factors as defined by similar GIM profiles, enlarged from the general picture shown in *A*. Asterisks indicate genes that are close to or overlap with the tested ORF deletions and for which the observed genetic interactions are correlated with the function of the other genes in the cluster. GO Term Finder [Boyle EI, *et al.* (2004) *Bioinformatics* 20:3710–3715] was used to assess the over-representation of a given cellular process or component in a cluster. (*C*) Same representation as in *B*, showing an example of clusters for ribosomal protein genes, formed mainly by common epistatic interactions with many query genes. While the profiles are similar, GIM correctly discriminated genes coding for proteins of the large and small ribosomal subunits. (*D*) Peroxisome related genes formed a well defined cluster despite the weak individual genetic interaction scores (maximum absolute log2(Q/R) values ranging from 1.01 to 1.56 for individual *pex* mutants).

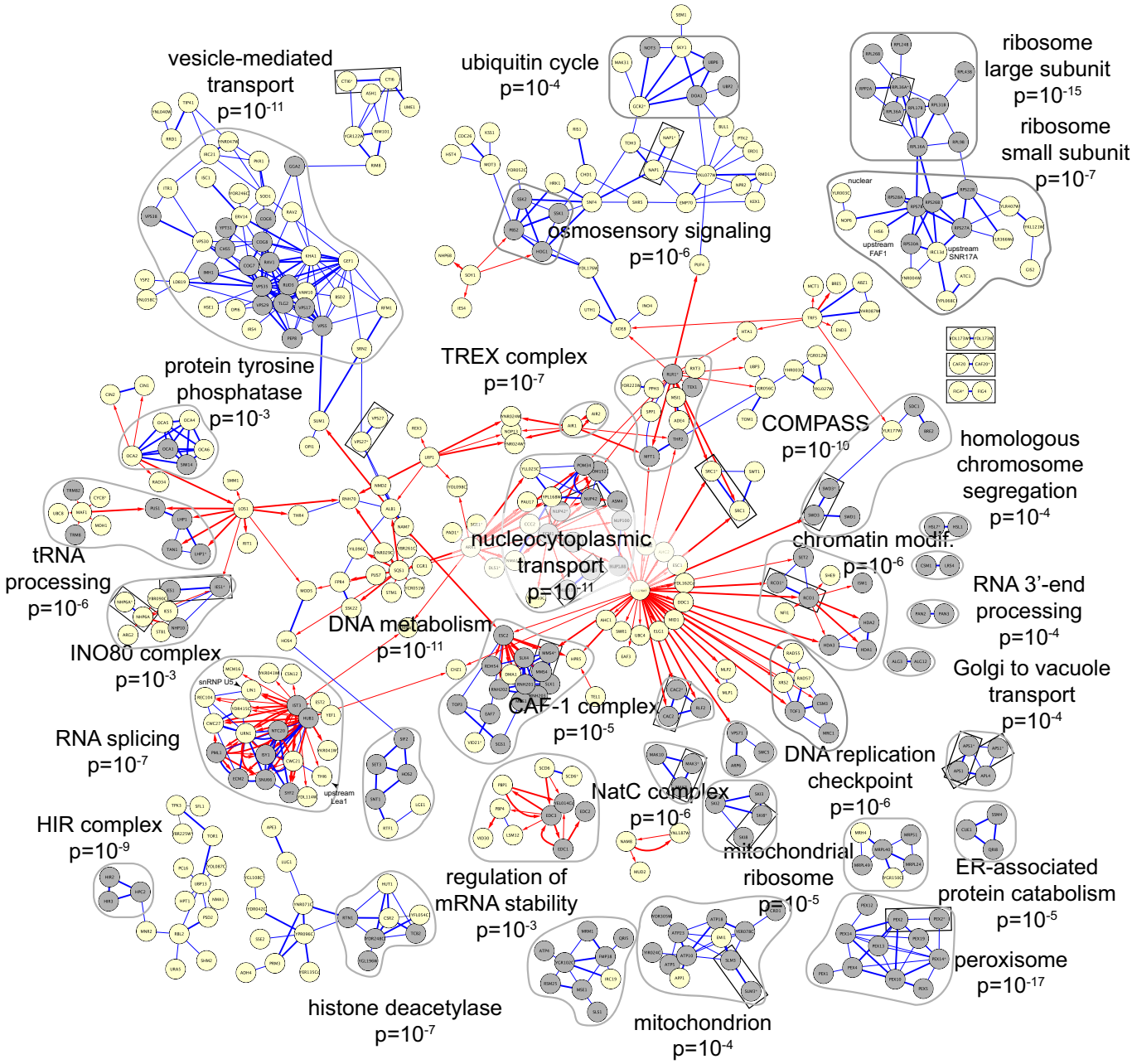

Fig. S7. A graphic representation of the network formed by the combined GIP and SGD links was obtained with Cytoscape [Shannon P, *et al.* (2003) *Genome Res* 13:2498–24504]. Several modules are defined by GO term Finder (associated *p* value indicated for each module) with annotated genes in gray circles. Red arrows and blue lines indicate synthetic growth defects and links generated from similar genetic interaction profiling (GIP) respectively.

Table S1. Proteins associated with Edc3-TAP and Scd6-TAP, and the two-hybrid partners of Scd6

ORF	Protein	No. of peptides
YBR094W	**Pby1**	26
YEL015W	**Edc3-CBP**	16
YDL160C	**Dhh1**	18
YEL015W	**Edc3***	9
YPR080W/YBR118W	Tef1/Tef2	17
YBR031W	Rpl4a	13
YNL118C	**Dcp2***	21
YLR340W	Rpp0	8
YOL149W	**Dcp1**	12

ORF	Protein	No. of peptides
P32503	**Virus L-A**	2 (MS/MS)
YPL106C	Sse1	11
YAL005C/YLL024C	Ssa1/Ssa2	16
YDL229W/YNL209W	Ssb1/Ssb2	13
YER165W	**Pab1**	10
YEL015W	**Edc3**	17
YDL160C	**Dhh1**	13
YAL038W	Cdc19	14
YPR129W	**Scd6-CBP**	10
YPR080W/YBR118W	Tef1/Tef2	14
YHR174W	Eno2	6
YOR063W	Rpl3	13
YBR031W/YDR012W	Rpl4a/Rpl4b	12
YPR129W	**Scd6***	5
YGR192C	Tdh3	9
YLR340W	Rpp0	8
YML063W	Rps1b	14
YFR031C-A/YIL018W	Rpl2a/Rpl2b	9
YNL178W	Rps3	6
YJR145C/YHR203C	Rps4a/Rps4b	6
YGR214W/YLR048W	Rps0a/Rps0b	6
YOL149W	**Dcp1**	18
YNL178W	Rps3	5

ORF	Protein	No. of clones
YNL118C	**Dcp2**	7 (7)
YEL015W	**Edc3**	3 (3)
YMR214W	Scj1	4 (1)
YHR186C	Kog1	1
YKR096W	Ykr096w	1
YPR184w	Gdb1	1
YDR206w	Ebs1	1
YLR086W	Smc4	1
YHL024W	Rim4	1
YAL026c	Drs2	1
YNR050c	Lys9	1
YLR067C	Pet309	1

Top and middle parts: MALDI-TOF-identified proteins are indicated, with factors considered to be specific in boldface. The asterisks indicate probable protein fragments. The number of peptides used for mass spectrometry identification is indicated. Bottom part: Preys selected in the Scd6 two-hybrid screen. The number of selected clones is indicated with the number of different fusions between parentheses. Proteins known to be involved in mRNA decapping are indicated in bold.

Table S2. Query mutant groups used in SGD score calculations

Group	Gene	ORF	Function
A	CWC21	YDR482C	Pre-mRNA splicing
	NAM8	YHR086W	Pre-mRNA splicing
	IST3	YIR005W	Pre-mRNA splicing
	PML1	YLR016C	Retention of unspliced pre-mRNAs
	PML39	YML107C	Retention of unspliced pre-mRNAs
	HUB1	YNR032C-A	Ubiquitin-like protein modifier
	URN1	YPR152C	Pre-mRNA splicing
B	PBP4	YDL053C	Lsm like protein
	EDC3	YEL015W	mRNA decapping
	EDC2	YER035W	mRNA decapping
	EDC1	YGL222C	mRNA decapping
	PBP1	YGR178C	Lsm like protein
	LSM12	YHR121W	Lsm like protein
	SCD6	YPR129W	Lsm like protein
C	NMD4	YLR363C	Nonsense-mediated mRNA decay
	NAM7	YMR080C	Nonsense mediated mRNA decay
D	AIR2	YDL175C	Nuclear mRNA surveillance
	LRP1	YHR081W	Nuclear exosome component
	AIR1	YIL079C	Nuclear mRNA surveillance
	YNL140C	YNL140C	Protein of unknown function
	TRF5	YNL299W	Poly(A) polymerase
	YNR024W	YNR024W	Nuclear exosome component
E	ESC2	YDR363W	Mating-type locus silencing
	RCO1	YMR075W	Histone deacetylase complex
D, E, F	NUP60	YAR002W	Nucleoporin
F	ECM1	YAL059W	Protein of unknown function
	ARX1	YDR101C	Pre-60S factor
	NUP42	YDR192C	Nucleoporin
	ALB1	YJL122W	Pre-60S factor
	SQS1	YNL224C	Pre-40S factor
G	MAF1	YDR005C	Negative regulator of RNA polymerase III
	LOS1	YKL205W	Nuclear export of pre-tRNA
H	RBG1	YAL036C	Protein of unknown function
	TMA46	YOR091W	Protein of unknown function
I	SOY1	YBR194W	Protein of unknown function
	NHP6A	YPR052C	Non-histone chromatin protein
Ungrouped	TEL1	YBL088C	Telomere length regulation
	CTH1	YDR151C	Protein of unknown function
	SUB1	YMR039C	Transcriptional coactivator
	OCA2	YNL056W	Protein of unknown function
	YNL187W	YNL187W	Protein of unknown function

The 41 query genes used in the GIM screens are listed with a summary of the function, when known. The genes were grouped depending on their function and the groups used in the calculation of the SGD specificity scores (see *SI Methods*). NUP60 was included into several groups.

Other Supporting Information Files

Dataset S1 (XLS)
Dataset S2 (XLS)
Dataset S3 (XLS)
Dataset S4 (XLS)