DISSECTING PATHWAYS WITH THE YEAST KNOCKOUT COLLECTION

by
Brian D. Peyser

A dissertation submitted to Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
March, 2007

Abstract

The yeast knockout collections provide opportunities to perform massively parallel phenotyping of deletion mutants for almost every yeast open reading frame. I used the knockout collection to screen for synthetic lethal partners, defined as alleles that cause lethality when combined but are nonlethal alone, with *CTF4* and *CTF18* and present the results in Chapter 2. I developed procedures for interpreting microarrays designed to compare changes in oligonucleotide TAGs specific to each knockout strain and present those methods in Chapters 3 and 4. These TAG microarrays allow thousands of experiments to screen for synthetic lethality among pairs of null alleles to be accomplished relatively quickly. In Chapter 4, I present 1410 novel predicted synthetic lethal interactions based on 707 currently completed screens. Interpretation of synthetic lethality is presented with a computational approach in Chapter 5, termed the congruence score. High congruence scores associate genes into common pathways, and I use the method to predict that *YLL049W* is a component of the dynein-dynactin nuclear orientation pathway. In Chapter 6, I propose a generalization of the congruence score to any phenotype, such as growth rate in the presence of various compounds, or even non-quantitative phenotypes such as cell morphology. This procedure connects genes based on similarity of multiple phenotypes using an application of information theory to produce a shared information score. Using gene ontology similarity, I show that high scores are associated with similarly annotated genes.

Advisor: Forrest A. Spencer PhD

Reader: Rafael A. Irizarry PhD

Preface

My interest in biology began with an early introduction to advanced placement biology as a freshman in high school, followed by dual enrollment in biotechnology as a senior. Thanks to my teacher Dr Susan Behel, I performed my first restriction digests and PCRs before high school graduation in 1993. The importance of these early biology courses might be exemplified by the thesis of one of my high-school friends, also a beneficiary of Dr Behel's instruction. Dr Noel Southall is a self-described "expert on water," which is accurate considering his biophysics dissertation on the hydrophobic effect.

This exposure to molecular biology led me to a position as a student research technician in Dr Ernest Hiebert's plant virology laboratory at the University of Florida Institute for Food and Agricultural Sciences. Dr Hiebert and his postdoctoral scientist Dr Ahmed Abouzid provided opportunities for me to learn various molecular biology techniques, and my summer work with another of Dr Hiebert's postdoctoral researchers, Dr Wayne Hunter, introduced me to electron and fluorescence microscopy.

The experience provided by my work in plant virology was a tremendous opportunity, and Dr George Agrios, the chairman of Plant Pathology, deserves great credit for the funding the department provided during my first year as a student assistant. When I met with my advisor, Dr F. William Zettler, during my first semester at the University of Florida and mentioned that I was interested in finding a job of that scope, he put me in touch with Dr Agrios, who was offering to fund research jobs for undergraduates. Dr Zettler was a great resource for me, and one of the reasons I chose

Plant Pathology as a second major in addition to Microbiology. With only 6 undergraduates in the program when I began, Plant Pathology was my true home department since Microbiology was full of approximately 100 times that many students, many of whom were in the pre-medicine track. While I experienced interesting coursework in the Microbiology department, Plant Pathology's Fifield Hall was where I learned truly applicable lessons.

Another important lesson from early in my undergraduate years was the ease with which seemingly insurmountable difficulties can be overcome when I have the support of my wife. Dr Rachel Wander-Peyser was my girlfriend at the time but she supported me utterly and continues to do so today. I experienced graduate school without her for one year at the University of Pennsylvania, and her absence was evidenced by the letter of academic probation that I received prior to leaving the program. I hope that my return to the University of Florida was helpful to her during veterinary school, but I know that she was essential to my success in graduate school at The Johns Hopkins University.

While my wife was becoming a veterinarian, I continued learning about biology as a senior lab tech and biological scientist in Dr Maureen Goodenow's pediatric HIV pathogenesis laboratory. My later interest in functional genomics began during my interactions with Dr Carter Coberley, who was a graduate student in Dr Goodenow's lab. Dr Coberley studied gene expression changes in cultured human macrophages infected with HIV using Affymetrix *GeneChips*. He and his wife, Dr Sadie Coberley, have been good friends and valuable advocates since then.

During my graduate school career I received significant help from many people,

including those acknowledged in following chapters, but some deserve mention here. My thesis committee members, Dr Geraldine Seydoux, Dr Susan Michaelis, Dr Neil Clarke, Dr David Cutler, and Dr Rafael Irizarry all provided helpful insights. Dr Cutler provided vital help in initiating the congruence score calculation described in Chapter 5, and he prevented huge wastes of time I would have otherwise endured with phenotypic similarity measures described in Chapter 6. Dr Irizarry was delightful to work with in the high-throughput synthetic lethality group, and he provided indispensable biostatistics expertise that made Chapters 3 and 4 possible.

The final committee member, my advisor Dr Forrest Spencer, deserves many thanks. I entered graduate school with more experience than most of my classmates, along with almost an expectation that graduate school would not end before some kind of screaming altercation with my mentor. In fact, Dr Spencer has been completely supportive and I have never been pushed to do undesired experiments or to follow questions that do not enthrall me. No amount of experience can guarantee a graduate student experience as enjoyable as mine has been.

Last, I must thank my son, Rohan Phineas Peyser. Rohan is almost one year old and is the best-behaved baby in the history of mankind. At times he allowed me to write this dissertation while my wife, Rachel, worked late, by playing quietly and looking at his books while the dogs played babysitter. I don't think he actually read any of his books on his own, but he does know the words "Mama," "Dada," and "uh-oh."

Table of Contents

List of Tables

Chapter 2 Tables

Chapter 3 Tables

Chapter 4 Tables

Chapter 5 Tables

Chapter 6 Tables

List of Figures

Chapter 2 Figures

Chapter 3 Figures

Chapter 4 Figures

Chapter 5 Figures

Chapter 6 Figures

Chapter 1.

Introduction

A significant advance in our ability to understand the biological pathways required for life was made by the introduction of the yeast knockout (YKO) collections (Giaever et al. 2002). These collections of *Saccharomyces cerevisiae* deletion mutants contain strains with defined null alleles for > 95% of predicted open reading frames (ORFs) in the yeast genome. Strains were created in haploid *MAT*a or *MAT*α backgrounds, *MAT*a/*MAT*α diploids (*yko*Δ/*yko*Δ), or as heterozygous diploids (*YKO*/*yko*Δ). Each mutant was created by homologous recombination with selectable marker KanMX4, consisting of a constitutive promoter from *Eremothecium gossypii* (*Ashbya gossypii*) fused to *nptI*, the kanamycin resistance gene.

The KanMX4 module was amplified in a polymerase chain reaction (PCR) using 74–75 bp oligonucleotide primers designed specifically for each deletion mutant. The primers contained, from 5′ to 3′, 18 or 19 nucleotides of homology to KanMX4 (U2 or D2), a 20 base TAG sequence particular to the targeted ORF, an 18 base universal priming site (U1 or D1), then 18 bases of sequence homologous to the 18 bp immediately up- or downstream of the predicted ORF. A second PCR was performed with the first product as template, using primers with homology to the 45 bp immediately up- or downstream of the ORF. This second PCR resulted in a construct with (from 5′ to 3′): 45 bp of homology to sequence immediately upstream of an ATG start codon, an 18 bp universal primer sequence (U1), a 20 bp "barcode" sequence (UPTAG) specific to the deletion mutant, the KanMX4 module, a 20 bp DNTAG, another 18 bp universal primer (D1), then 45 bp of homology to the genomic sequence immediately downstream of the stop codon (Winzeler and Shoemaker et al. 1999; Giaever et al. 2002; http://www-

). The end result was a collection of YKO strains with specific ORFs precisely deleted, each marked with oligonucleotide TAGs and the KanMX4 G418-resistance cassette.

Fewer than 20% of these deleted ORFs are required for survival on rich glucose medium. While the essential genes provide important functions, a large majority of ORFs encode proteins that are not required for growth under any standard laboratory condition. This suggests that many yeast genes are evolutionarily retained for their ability to buffer genetic variation in other genes (Hartman et al. 2001).

Alleles that are lethal in combination but individually nonessential are termed "synthetic lethal" (Dobzhansky 1946). Sometimes this occurs when a point mutation in a component of an essential protein complex is combined with a point mutation from another component of that same complex (Potenza et al. 1992; Appling 1999). The combination of point mutations results in failure to organize the essential complex and therefore, death of the cell. When considering null mutations, dispensable genes may be buffered by another pathway, which results in lethality when components of the buffering pathway are removed (see Guarente 1993).

This genetic buffering due to parallel pathways is the current focus of study due to introduction of the YKO collections. Multiple studies have been performed using the entire set of viable deletion mutants to screen for synthetic lethality. Some of these studies use arrays of individual YKOs combined with a query mutation and score genetic interaction based on direct measurement of colony size (Tong et al. 2001; Tong et al. 2004; Tong and Boone 2006). This method is termed synthetic genetic array (SGA) or

epistatic miniarray profile (E-MAP) for a subset approach to SGA (Schuldiner et al. 2005; Collins et al. 2006; Schuldiner et al. 2006; Collins et al. 2007). Other approaches estimate growth defects by microarray hybridization of the oligonucleotide TAGs from samples grown competitively in a pool, known as synthetic lethality analyzed by microarray (SLAM) (Ooi et al. 2003), or diploid-based SLAM (dSLAM) (Pan et al. 2004; Pan et al. 2007).

The rapidity of SLAM provides an opportunity to create all ~25 million double mutant combinations and probe their growth rates using microarrays. We have begun this process and have currently completed about 1/7 of the genome (see Chapter 4). With genome-wide synthetic lethality information, pathways responsible for growth can be deduced based on the pattern of interactions (see Chapter 5). The null alleles provided by the YKO collections exhibit synthetic lethality between genes in separate, related, parallel pathways (Guarente 1993).

Other experiments have assigned genes to common pathways based on similarity of expression patterns (see Hughes and Marton et al. 2000 for a landmark study of this kind). However, this approach simply finds common transcription modules, which is not the same as common pathways. Genes that function in the same pathway often share transcription control mechanisms, for example, the *lac* operon in *Escherichia coli*. Comparison with data from functional profiling of the YKO collections shows that genes up-regulated in response to specific treatments are rarely included among those genes required for optimal growth in those same conditions (Giaever et al. 2002; Birrell et al. 2002). This argues for use of YKO phenotypes rather than expression patterns to assign

pathway membership, and in Chapter 6 I present a calculation that uses a collection of these phenotypes to align genes into common pathways.

References

Appling DR. 1999. Genetic approaches to the study of protein-protein interactions. *Methods* 19(2):338–49.

Birrell GW, Brown JA, Wu HI, Giaever G, Chu AM, Davis RW, Brown JM. 2002. Transcriptional response of Saccharomyces cerevisiae to DNA-damaging agents does not identify the genes that protect against these agents. *Proc Natl Acad Sci USA* 99(13):8778–83.

Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, Schuldiner M, Gebbia M, Recht J, Shales M, et al. 2007. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* [Epub ahead of print 2007-02-21].

Collins SR, Schuldiner M, Krogan NJ, Weissman JS. 2006. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol* 7(7):R63.

Dobzhansky T. 1946. Genetics of natural populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura*. *Genetics* 31(3):269–90.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418(6896):387–91.

Guarente L. 1993. Synthetic enhancement in gene interaction: a genetic tool come of

age. *Trends Genet* 9(10):362–6.

Hartman JL 4th, Garvik B, Hartwell L. 2001. Principles for the buffering of genetic variation. *Science* 291(5506):1001–4.

Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai H, He YD, et al. 2000. Functional discovery via a compendium of expression profiles. *Cell* 102(1):109–26.

Ooi SL, Shoemaker DD, Boeke JD. 2003. DNA helicase interaction network defined using synthetic lethality analyzed by microarray. *Nat Genet* 35(3):277–86.

Pan X, Yuan DS, Ooi SL, Wang X, Sookhai-Mahadeo S, Meluh P, Boeke JD 2007. dSLAM analysis of genome-wide genetic interactions in *Saccharomyces cerevisiae*. *Methods* 41(2):206–21.

Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD. 2004. A robust toolkit for functional profiling of the yeast genome. *Mol Cell* 16(3):487–96.

Potenza M, Bowser R, Muller H, Novick P. 1992. *SEC6* encodes an 85 kDa soluble protein required for exocytosis in yeast. *Yeast* 8(7):549–58.

Schuldiner M, Collins SR, Thompson NJ, Denic V, Bhamidipati A, Punna T, Ihmels J, Andrews B, Boone C, Greenblatt JF, et al. 2005. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* 123(3):507–19.

Schuldiner M, Collins SR, Weissman JS, Krogan NJ. 2006. Quantitative genetic

analysis in *Saccharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. *Methods* 40(4):344–52.

Tong AH, Boone C. 2006. Synthetic genetic array analysis in *Saccharomyces cerevisiae*. *Methods Mol Biol* 313:171–92.

Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, et al. 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294(5550):2364–8.

Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, et al. 2004. Global mapping of the yeast genetic interaction network. *Science* 303(5659):808–13.

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, et al. 1999. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901–6.

Chapter 2.

Examining the functions of CTF4 and CTF18 with global synthetic

lethality screens

This work contributed to the publication:

Introduction

In budding yeast, *CTF4* and *CTF18* are required for robust sister chromatid cohesion (Hanna et al. 2001). This cohesion is carried out by the essential Cohesin complex, consisting of Scc1p, Scc3p, Smc1p, and Smc3p (reviewed in Nasmyth 2001). Sister chromatids are produced and cohesion is established during S phase (Skibbens et al., 1999; Toth et al., 1999). The association of sister chromatids must be maintained until anaphase, when sister chromatids are split resulting in one of each chromosome in the daughter nucleus.

When cells are arrested at metaphase by depolymerization of microtubules, cohesion failure occurs rarely in wild-type yeast, but *ctf4Δ* and *ctf18Δ* mutants exhibit failure in ~30% of cells, a significant increase (Hanna et al. 2001; Mayer et al. 2001). When *ctf4Δ* and *ctf18Δ* alleles are combined in the same cell, yeast are unable to survive (Miles and Formosa 1992; Formosa and Nittis 1999). This genetic interaction is termed synthetic lethality, and provides information about what functions are required in the absence of some other nonessential function. In order to more fully understand the roles of *CTF4* and *CTF18* in contributing to sister chromatid cohesion, we searched for other synthetic genetic interactions using a genome-wide screen.

Materials and Methods

Random Spore Analysis

A set of candidate synthetic lethal interactions produced by J. Hanna using a haploid synthetic lethality analyzed by microarray (SLAM) method (Ooi et al. 2003) prior to my

involvement was tested against both *ctf4Δ* and *ctf18Δ* using random spore analysis.

Candidate *MAT*a deletion mutants (*ykoΔ::KanMX*, *yko* represents y̲east k̲nock̲out) from

the YKO collection (Giaever et al. 2002) were mated to strains YJH96 (*MATα*

ctf4Δ::NatMX can1::MFA1pr-HIS3) and YJH97 (*MATα ctf18Δ::NatMX can1::MFA1pr-*

HIS3) on solid yeast extract, peptone, dextrose media (YPD), and diploids were selected

on solid YPD +200 μg/ml G418 (Cellgro, Herndon VA) +100 μg/ml clonNAT (Hans-

Knöll Institute für Naturstoff-Forschung, Jena Germany). The resulting heterozygous

diploids were replica-plated onto solid sporulation media and grown at 25 °C for 5 d.

Following sporulation, a swatch of cells was transferred into 500 μl of sterile dH$_2$O and

briefly sonicated to break apart cell clumps, then 10, 20, or 40 μl were plated on solid

haploid selection media (synthetic complete −His −Arg +50 μg/ml canavanine [Sigma-

Aldrich, St Louis MO]) to select all *MAT*a spores, +200 μg/ml G418 to select single

mutants, or +200 μg/ml G418 +100 μg/ml clonNAT to select double mutants,

respectively. Monosodium glutamic acid (1 g/l) replaced ammonium sulfate in synthetic

media when G418 or clonNAT selections were applied. Growth of double mutant

colonies was compared to single mutants after 42 h at 30 °C, and double mutants that

failed to grow or grew more slowly were scored as synthetic lethal or fitness defect.

Synthetic Lethality Screen

Screens for synthetic lethal mutants were repeated using an early diploid-based

SLAM (dSLAM) method (Pan 2007). Heterozygous diploid yeast knockout (YKO)

strains (Research Genetics) were pooled, then modified by introducing *can1Δ::MFA1pr-*

HIS3, a *MAT*a haploid selection marker (Tong et al. 2001), by transformation *en masse*

10

with the targeting construct (gift of X. Pan, Baltimore). This haploid-convertible pool was transformed with a PCR construct consisting of ~2 kb genomic sequence surrounding *CTF4* on either side of the *NatMX* nourseothricin resistance cassette. In parallel, transformations were performed with *ura3Δ::NatMX* to serve as a control. Approximately 5×10^5 transformants were selected on solid YPD + 200 μg/ml G418 +100 μg/ml clonNAT. These double mutant heterozygous diploids were scraped into a pool of ~2×10^8 cells and cultured in liquid sporulation media for 7 d. The cells were checked for asci to verify sporulation, then cultured on solid SC −His −Arg +50 μg/ml canavanine +G418 +clonNAT to select for ~10^6 *MAT*a haploid double mutants. Each experiment was performed in duplicate.

Tag Microarray Hybridization

Haploid double mutants were collected and ~2×10^8 cells were used to prepare genomic DNA. Biotin-labeled UPTAGs and DNTAGs were generated from each sample using ~200 ng genomic DNA as template in a PCR with biotinylated primers as previously described (Giaever et al. 2002). PCR products were separated from unincorporated primers using Microcon YM-10 columns. These labeled UPTAGs and DNTAGs were combined and hybridized to Tag3 arrays (Affymetrix) as described (Giaever et al. 2002).

Microarray Analysis

Signal intensities were read from Affymetrix ".cel" files into R (Ihaka and Gentleman 1996) and analyzed using the Bioconductor *affy* package (Gautier et al. 2004) and custom script. Perfect match and complement perfect match signal values were

averaged, then $\log_2$-transformed. Mismatch probes were not used. TAGs from all essential genes and TAGs that had signal intensities lower than 97% of all essential genes on both URA3 control chips were removed from the data separately for UPTAG and DNTAG. Remaining UP- and DNTAGs were then quantile normalized (Bolstad et al. 2003). All pairwise $\log_2$ ratios were generated between experiment and control chips ($\log_2$[*ura3*a/*ctf4*a], $\log_2$[*ura3*a/*ctf4*b], $\log_2$[*ura3*b/*ctf4*a], and $\log_2$[*ura3*b/*ctf4*b]), and then averaged. For comparisons of known synthetic lethal/fitness defect results with genes tested with no interaction, the UP- and DNTAG $\log_2$ ratios were averaged. When one of the TAGs was filtered, the other value was used. For candidate interactions the larger of UP- or DNTAG $\log_2$ ratio was used, and any YKO with a log ratio ≥ 1.1 and $\left| \log_2 (ura3\,a\,/\,ura3\,b) \right| \leq 0.75$ was chosen.

Results

The combined function of *CTF4* and *CTF18* is not optional for viability in yeast. This suggests that each gene provides some similar function in parallel, either one of which is nonessential. Using random spore analysis, we examined synthetic lethal or fitness defect interactions in both *ctf4*Δ and *ctf18*Δ backgrounds to examine the requirements of each mutant. Selection of *MAT*a haploid spores with varying marker requirements was made possible by *MFA1pr-HIS3* and *can1*[R] (Tong et al. 2001). Random spore analysis displayed differences in growth rate between single and double mutants as smaller colonies under double selection. We tested 84 potential interactions with both *ctf4*Δ and *ctf18*Δ and an additional 18 interactions with *ctf4*Δ alone. Interactions were scored without knowledge of the identity of each mutant, and given a score of 0 (no

12

interaction), 1 (slight fitness defect), 2 (fitness defect), or 3 (lethal). Figure 1 shows examples for each level of fitness defect. This random spore analysis revealed synthetic fitness defects with *ctf4Δ* for 31 of the 102 genes (Table 1). Of these, only *ctf18Δ* (*chl12Δ*) had been previously reported (Formosa and Nittis 1999). We also found 13 genetic interactions with *ctf18Δ*. Among the 15 interactions exhibited by *ctf4Δ* for which we have information on *ctf18Δ*, 12 also interacted with *ctf18Δ*. The shared interactions suggest that *CTF4* and *CTF18* provide similar functions—loss of either one is lethal in cells lacking *CLB2*, *FUR4*, or *HPR5*.

We performed a microarray-based synthetic lethal screen (see Ooi et al. 2003 and Methods) to search for additional nonessential genes that require *CTF4*. This technique examines the relative abundance of "barcode" TAGs uniquely marking each deletion mutant (Shoemaker et al. 1996) in two populations of pooled yeast knockout strains (Giaever et al. 2002). A pool of 5916 *YKO/ykoΔ::KanMX* mutants was transformed *en masse* with a *ctf4Δ::NatMX* deletion construct, to generate a heterozygous pool of double mutants. Replacement of *ura3Δ0* with *ura3Δ::NatMX* was performed in parallel to serve as a control. Haploid *MAT*a double mutants were selected from each pool and the relative representation of each *ykoΔ::KanMX* mutant was compared between *ura3Δ::NatMX* and *ctf4Δ::NatMX* (control:experiment). For those mutants unable to grow in the presence of *ctf4Δ*, we expect to find a large control:experiment ratio.

Microarray results for *ctf4Δ* suggested synthetic lethal interactions for many of the known genes as well as some new potential interactions (see Table 2, pursued in Warren et al. 2004). The average $\log_2$ ratio was significantly correlated for the known mutants

(Figure 2). The Spearman correlation coefficient was 0.57 for mean $\log_2$ ratio versus interaction score, with $P < 0.001$. The $\log_2$ ratios were significantly higher for known synthetic lethal compared to known healthy ($P = 5 \times 10^{-7}$), for known synthetic fitness defect versus known healthy ($P = 1 \times 10^{-7}$), and synthetic lethal pairs displayed higher ratios than synthetic fitness defect pairs ($P = 0.03$) by the Wilcoxon rank sum test. The median and interquartile ranges were 1.01 and 0.71 to 1.55, 0.57 and 0.26 to 0.74, −0.03 and −0.24 to 0.17 for known synthetic lethal, synthetic fitness defect, and known healthy, respectively.

Discussion

We discovered 30 novel synthetic genetic interactions with *ctf4Δ*. These interactions provided interesting targets for analysis of sister chromatid cohesion (Warren et al. 2004). One interesting result is the lethal phenotype observed in *ctf4Δ fur4Δ* double mutants (Table 1). *FUR4* encodes uracil permease, providing a route for entry of uracil in the culture medium into the cell. The *ctf4Δ* strains we used were Ura⁺, without which *fur4Δ* cells would not survive on synthetic media. However, even in a Ura⁺ cell, loss of the ability to take up uracil could affect the levels of UMP, and therefore UTP and CTP, due to decreased flux through the salvage pathways of pyrimidine ribonucleotides. Rather than generating UMP from uracil through the salvage pathway, *fur4* mutants must utilize the de novo biosynthesis pathway. This could impact levels of PRPP and reduce the availability of purines and deoxyribonucleotides as well. Since *CTF4* is involved in DNA replication, it is possible that stress to the availability of deoxyribonucleotides during DNA synthesis may cause DNA damage that requires *CTF4* for proper resolution.

14

The importance of *CTF4* in relation to DNA damage is reinforced by the synthetic lethal phenotypes with *hpr5Δ* and *mrc1Δ*. *HPR5* (*RADH/SRS2*) is a helicase involved in DNA repair (Aboussekhra et al. 1989; Rong and Klein 1993), and *MRC1* activates the S-phase checkpoint in response to DNA damage (Alcasabas et al. 2001). These interactions suggest that *CTF4* can provide an alternative DNA damage response, or that *CTF4* is required to properly replicate chromosomes in the face of increased DNA damage.

Acknowledgments

References

Bolstad BM, Irizarry RA, Astrand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19(2):185–93.

Formosa T, Nittis T. 1999. Dna2 mutants reveal interactions with Dna polymerase alpha and Ctf4, a Pol alpha accessory factor, and show that full Dna2 helicase activity is not essential for growth. *Genetics* 151(4):1459–70.

Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. affy—analysis of *Affymetrix GeneChip* data at the probe level. *Bioinformatics* 20(3):307–15.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al. 2002. Functional profiling of the *Saccharomyces*

cerevisiae genome. *Nature* 418(6896):387–91.

Hanna JS, Kroll ES, Lundblad V, Spencer FA. 2001. *Saccharomyces cerevisiae CTF18* and *CTF4* are required for sister chromatid cohesion. *Mol Cell Biol* 21(9):3144–58.

Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. *J Comput Graph Stat* 5(3):299–314.

Mayer ML, Gygi SP, Aebersold R, Hieter P. 2001. Identification of RFC(Ctf18p, Ctf8p, Dcc1p): an alternative RFC complex required for sister chromatid cohesion in *S. cerevisiae*. *Mol Cell* 7(5):959–70.

Miles J, Formosa T. 1992. Evidence that *POB1*, a *Saccharomyces cerevisiae* protein that binds to DNA polymerase alpha, acts in DNA metabolism in vivo. *Mol Cell Biol* 12(12):5724–35.

Nasmyth K. 2001. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev Genet* 35:673–745.

Ooi SL, Shoemaker DD, Boeke JD. 2003. DNA helicase interaction network defined using synthetic lethality analyzed by microarray. *Nat Genet* 35(3):277–86.

Shoemaker DD, Lashkari DA, Morris D, Mittmann M, Davis RW. 1996. Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* 14(4):450–6.

Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, et al. 2001. Systematic genetic analysis with

ordered arrays of yeast deletion mutants. *Science* 294(5550):2364–8.

Warren CD, Eckley DM, Lee MS, Hanna JS, Hughes A, Peyser B, Jie C, Irizarry R, Spencer FA. 2004. S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion. *Mol Biol Cell* 15(4):1724–35.
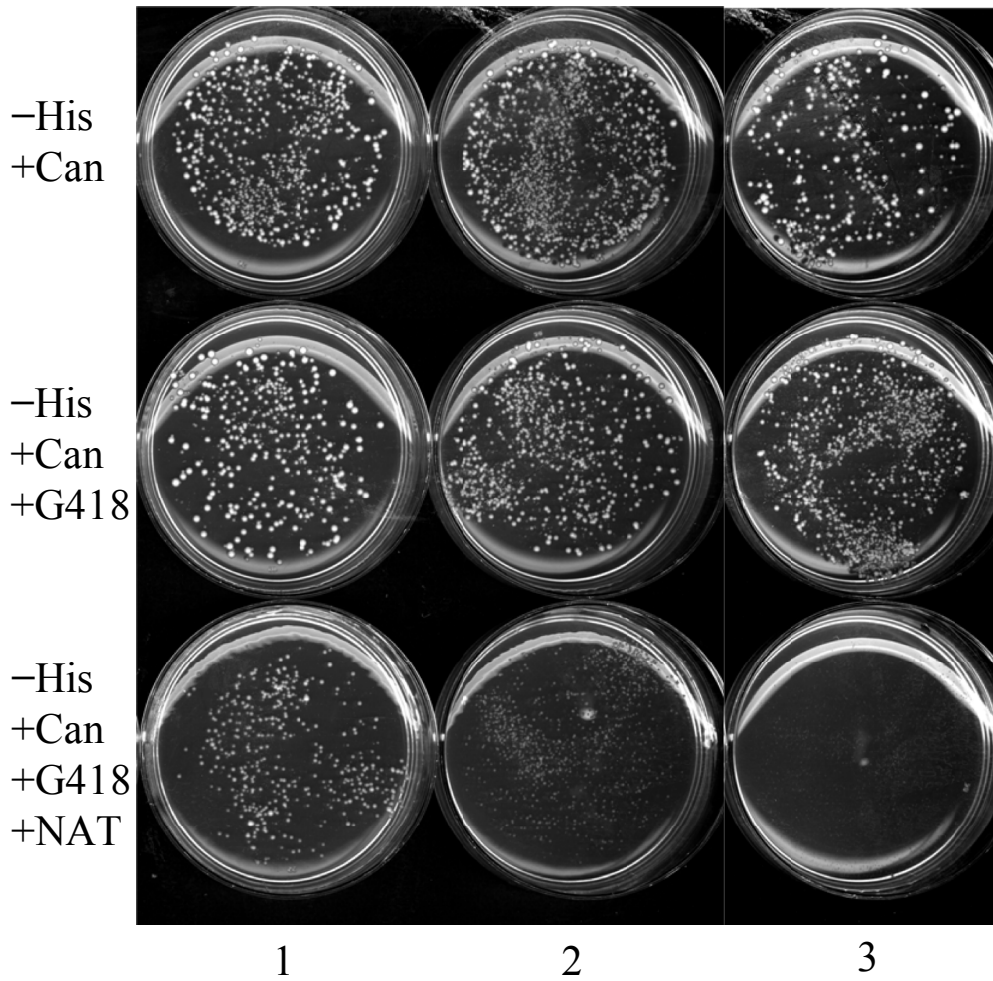
Figure 1. Examples of synthetic genetic interactions. Spores were selected on −His +Canavinine (all *MAT*a haploids), −His +Can +G418 (*MAT*a single mutants), and −His +Can +G418 +NAT (*MAT*a double mutants). Synthetic genetic interactions were scored as: 1 = slight fitness defect, 2 = fitness defect, 3 = lethal.

CTF4 Microarray Results
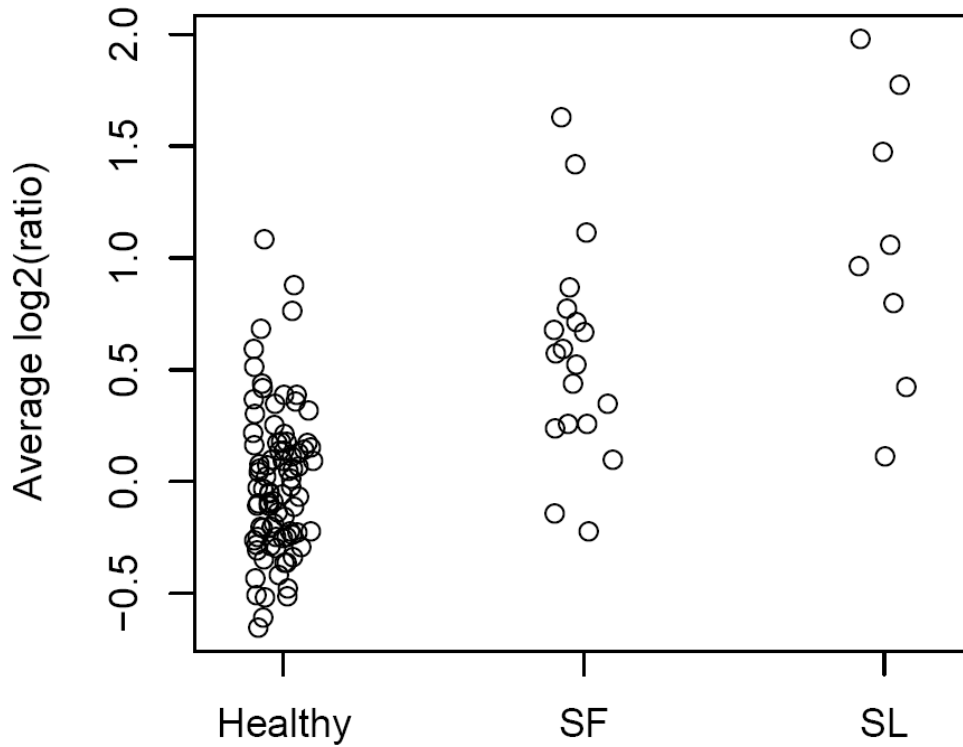
Figure 2. Microarray results. The average *ctf4*Δ $\log_2$ ratio from UPTAG and DNTAG for four comparisons (*ura3*a/*ctf4*a, *ura3*b/*ctf4*a, *ura3*a/*ctf4*b, and *ura3*b/*ctf4*b) was compared to known phenotypes. Healthy, genes that were tested with *ctf4*Δ and did not exhibit a synthetic genetic interaction; SF, genes that exhibited a synthetic fitness defect with *ctf4*Δ; SL, genes verified as lethal in combination with *ctf4*Δ.

Table 1. Synthetic genetic interactions. Double mutants for each listed gene with *ctf4Δ* and *ctf18Δ* were assayed for synthetic lethality by random spore analysis. Mutants were scored for growth: 0, no fitness defect; 1, slight synthetic fitness defect; 2, synthetic fitness defect; 3, synthetic lethal. Strains not tested with *ctf18Δ*, nd.

Gene	CTF4	CTF18	Gene	CTF4	CTF18	Gene	CTF4	CTF18
CLB2	3	3	ATG18	0	0	SLA1	0	0
FUR4	3	3	BUD17	0	0	SMM1	0	0
HPR5	3	3	BUL2	0	0	SNF11	0	0
MRC1	3	2	CHS5	0	0	SPO22	0	0
CTF18	3	nd	CTF3	0	0	SPT8	0	0
DCC1	3	nd	CTP1	0	0	SRV2	0	0
KAR3	3	nd	CYS3	0	0	SSK1	0	0
MSS18	3	nd	DOT1	0	0	STE50	0	0
CSM3	2	3	ECM34	0	0	SWP82	0	0
LTE1	2	3	EPS1	0	0	TRM10	0	0
TOF1	2	3	FIS1	0	0	WHI2	0	0
MAD1	2	2	FMP21	0	0	YBL012C	0	0
RIM8	2	2	GFD2	0	0	YBL094C	0	0
DOA1	2	1	GNP1	0	0	YBR284W	0	0
CAC2	2	0	HAL5	0	0	YCL033C	0	0
UBP3	2	0	HNM1	0	0	YCR087W	0	0
CTF19	2	nd	IDH1	0	0	YDR042C	0	0
ELM1	2	nd	IMG2	0	0	YHL029C	0	0
MCM21	2	nd	IRA2	0	0	YHR112C	0	0
MDM39	2	nd	MIG2	0	0	YIL077C	0	0
RAD61	2	nd	MNN11	0	0	YJL131C	0	0
RMD7	2	nd	NAS2	0	0	YKL023W	0	0
SWM1	2	nd	NPR2	0	0	YLL007C	0	0
YJR018W	2	nd	PCL7	0	0	YLR236C	0	0
CIN8	1	3	PDE2	0	0	YLR414C	0	0
MAD2	1	2	PDR18	0	0	YMR084W	0	0
MCM16	1	0	PKH2	0	0	YNR021W	0	0
BFA1	1	nd	PMS1	0	0	YNR047W	0	0
BUB2	1	nd	RMD11	0	0	YNR068C	0	0
CHL4	1	nd	RTG1	0	0	YPR170C	0	0
MID1	1	nd	RVS161	0	0	YPT6	0	0
SPT3	0	2	SCY1	0	0	YSA1	0	0
ACH1	0	0	SDH4	0	0	ARD1	0	nd
ARP6	0	0	SIW14	0	0	COG6	0	nd

Table 2. Microarray *ctf4Δ* synthetic lethal candidates. YKOs with average UP- or DOWNTAG log$_2$ ratio ≥ 1.1 and control:control log$_2$ ratio ≤ 0.75. YKOs already tested are marked under "Known" as SL, synthetic lethal; SF, synthetic fitness defect.

ORF	Gene	Process	Avg log$_2$ ratio UPTAG	Avg log$_2$ ratio DOWNTAG	Known
YPR134W	*MSS18*	mRNA splicing		2.19	
YPR119W	*CLB2*	G2/M transition of mitotic cell cycle	1.98		SL
YDR318W	*MCM21*	chromosome segregation		1.94	
YPR133W-A	*TOM5*	mitochondrial translocation	1.52	2.24	
YCL016C	*DCC1*	sister chromatid cohesion	2.29	1.26	SL
YHR013C	*ARD1*	protein amino acid acetylation	1.31	2.22	
YDR260C	*SWM1*	spore wall assembly		1.70	
YER083C	*RMD7*	cell wall organization and biogenesis	2.42	0.96	
YPR141C	*KAR3*	meiosis, mitosis		1.65	
YAL024C	*LTE1*	exit from mitosis	1.24	2.02	SF
YCL060C	*MRC1*	DNA replication checkpoint		1.47	SL
YJL030W	*MAD2*	mitotic spindle checkpoint	1.28	1.56	SF
YPL018W	*CTF19*	chromosome segregation	1.12	1.63	
YER014C-A	*BUD25*	bud site selection	1.29	1.21	
YJR018W			1.24		
YNL291C	*MID1*	calcium ion transport	0.97	1.32	
YGL045W				1.12	
YNL273W	*TOF1*	DNA topological change	1.24	0.98	SF
YDR410C	*STE14*	peptide pheromone maturation		1.10	
YJL092W	*HPR5*	DNA repair, NHEJ	1.74	0.38	SL
YDR254W	*CHL4*	chromosome segregation	0.92	1.14	
YDR014W	*RAD61*		1.19	0.80	
YMR190C	*SGS1*	chromosome segregation	0.72	1.19	
YMR055C	*BUB2*	mitotic spindle checkpoint	1.46	0.33	
YMR048W	*CSM3*	meiotic chromosome segregation	1.65	0.09	SF
YNL041C	*COG6*	intra Golgi transport	0.39	1.34	
YKL048C	*ELM1*	axial budding	0.51	1.21	
YDR114C			0.29	1.41	
YCL061C	*MRC1*	DNA replication checkpoint	1.24	0.37	SL
YGL020C	*MDM39*	mitochondrion organization and biogenesis	0.23	1.36	
YJR053W	*BFA1*	mitotic spindle checkpoint	0.21	1.19	

Chapter 3.

Improved statistical analysis of budding yeast TAG microarrays revealed by defined spike-in pools

Brian D. Peyser, Rafael A. Irizarry, Carol W. Tiffany, Ou Chen, Daniel S. Yuan, Jef D. Boeke and Forrest A. Spencer

Abstract

Saccharomyces cerevisiae knockout collection TAG microarrays are an emergent platform for rapid, genome-wide functional characterization of yeast genes. TAG arrays report abundance of unique oligonucleotide 'TAG' sequences incorporated into each deletion mutation of the yeast knockout collection, allowing measurement of relative strain representation across experimental conditions for all knockout mutants simultaneously. One application of TAG arrays is to perform genome-wide synthetic lethality screens, known as synthetic lethality analyzed by microarray (SLAM). We designed a fully defined spike-in pool to resemble typical SLAM experiments and performed TAG microarray hybridizations. We describe a method for analyzing two-color array data to efficiently measure the differential knockout strain representation across two experimental conditions, and use the spike-in pool to show that the sensitivity and specificity of this method exceed typical current approaches.

Introduction

Introduction of the yeast knockout collections, containing arrayed strains harboring deletion mutations for > 95% of predicted open reading frames (ORF), allows systematic genome-wide screens for various phenotypes to be readily accomplished (Winzeler and Shoemaker et al. 1999; Tong et al. 2001; Giaever et al. 2002). One aspect of the knockout collections that facilitates rapid screens is the pair of unique 20 nucleotide TAGs within each deletion mutation (Shoemaker et al. 1996). A gene in each yeast knockout strain (YKO) is replaced with a selectable marker flanked by two TAGs, termed UPTAG and

DNTAG. All UPTAGS or all DNTAGs in a sample can be amplified in a PCR reaction using universal primers. Individual YKO representation is subsequently interrogated by hybridizing labeled TAGs to microarrays and observing changes in signal intensity between experimental conditions. This approach has been applied to genome-wide screens for mutant phenotypes (Ooi et al. 2001; Giaever et al. 2002), synthetic genetic interactions (Ooi et al. 2003; Warren et al. 2004; Lee and Spencer, 2004; Arevalo-Rodriguez et al. 2004; Pan et al. 2004) and synthetic-chemical-genetic interactions (Pan et al. 2004). TAG microarray approaches are rapid and comprehensive, but systematic optimization of analysis methods is lacking. We address this need here.

The most common application of TAG arrays is comparison of YKO representation in two samples. Typically, samples are co-hybridized on one array using complementary fluorescent labels (Cy5 and Cy3). Various general approaches for two-color arrays have been proposed for quality assessment, background adjustment (Kooperberg et al. 2002; Yang et al. 2002) and normalization (Kerr et al. 2000; Dudoit et al. 2002; Huber et al. 2002). However, in analysis of TAG arrays, each YKO has UPTAG and DNTAG probes; four measurements corresponding to each YKO are obtained. Finding the best way to summarize this information in one quantity reflecting differential YKO representation is not trivial (Irizarry et al. 2003). Here we demonstrate the utility of a spike-in experiment by evaluating a simple quality assessment procedure and a novel strategy for combining the UPTAG and DNTAG information in a way that is robust to problematic TAGs. Our results show that implementing these two data procedures can greatly improve the utility of TAG arrays.

24

Materials and methods

Preparation of spike-in pools. Heterozygous YKOs (Research Genetics) were grown on solid media and combined, with the exception of YKOs from plate 259 which were separately mixed in subpools, as shown in Figure 1, before incorporation into pool A or B at appropriate representation levels. Genomic DNA was extracted from samples of each spike-in pool using the Masterpure Yeast DNA kit (Epicentre). Pool A and B TAGs were labeled with Cy3 and Cy5, respectively, and TAG microarrays were hybridized, washed and scanned as described (Yuan et al. 2005).

Analyses were performed using custom scripts written in R, an open-source statistical language (Ihaka and Gentleman 1996). GenePix local background intensities were not used for correction because, as suggested by Yang et al. (2002), subtracting these severely increases noise (data not shown). Normalization was performed using a procedure similar to the one previously proposed (Dudoit et al. 2002). Alternate normalization methods did not impact results (data not shown).

The GEO accession number for microarray data is GSE2832. Data and code necessary to reproduce all the results and figures are available upon request.

Results

To evaluate statistical procedures for TAG microarray data, we tailored defined spike-in pools to resemble expected results in a typical synthetic lethality analyzed by microarray (SLAM) experiment (Ooi et al. 2003). Synthetic lethality is defined as inviability of cells containing two mutations which are individually not required for

growth. In a SLAM experiment, viable YKOs in pooled form are compared under two

conditions: absence versus presence of a specific second mutation (the 'query'). The

average number of genetic interactions expected in a genome-wide screen has been

estimated to be ~35, although several query mutations with interactions exceeding 100

have been analyzed (Pan et al. 2004; Tong et al. 2004). Therefore, we designed a pair of

pools ('A' and 'B') with 5758 YKOs at equivalent representation, and a set of 94 YKOs

with known differential representation ranging from $1:2^{\frac{1}{3}}$ to $1:2^5$ and 1:infinity (Figure 1

and Supplementary Table 1). Additionally, certain YKOs grow slowly, and

representation of these in the control SLAM sample is expected to be lower than YKOs

with wild-type growth rate. To examine our ability to address these mutants, we designed

three representation levels in the control (B) pool: high (about equal to all other strains),

medium (8-fold dilute) and low (64-fold dilute). TAGs from pools A and B were

amplified with Cy3- and Cy5-labeled primers, respectively. These samples were mixed at

equal ratio, such that most TAGs should exhibit equal hybridization, while Cy5:Cy3

ratios that deviate from one are expected for the few differentially represented TAGs.

This design allows discovery of the best method to produce a measure of differential

representation from hybridization results.

Before addressing differential representation, we document the utility of two

filtering steps in data pre-processing. First, we noted TAG-specific hybridization artifacts

evident in self-self hybridizations performed to examine the noise distribution. Pool A

DNA served as template for preparation of labeled TAGs with both Cy5 and Cy3. Thus,

all TAGs were present at equal amounts between channels. Figure 2a shows a scatterplot

of normalized $\log_2$ Cy5:Cy3 ratios for corresponding UPTAG and DNTAG probes.

Because all these values should be zero plus measurement error, we expect these to be uncorrelated and centered at zero. Figure 2a confirms this except for a few YKOs with extreme values for one TAG. These outliers may have a negative impact on specificity. They are likely to be due to individual tag templates that enter the labeling PCR as contaminants, which are detectable even at very low levels (Yuan et al. 2005).

We determined that these artifacts are consistent across experiments performed with a single batch of labeled primer, but not between different primer batches (Supplementary Figure 1). To create a useful filter, we assumed that the data follow a bivariate normal distribution and defined outliers as TAGs with log ratios three SDs away from zero, using a robust estimate of the SD. If the log ratio data follow a normal distribution, excluding outliers, we expect to inappropriately remove only ~0.5% of the data (32 TAGs). We applied this filter (Figure 2a, red lines) independently to UP- and DNTAGs. Fortunately, the YKOs were designed with two TAGs per gene (except for 192 strains lacking DNTAGs), greatly improving chances that at least one TAG performs adequately. Because non-outlier UP/DNTAGs appear to provide independent measurements, the chance of inappropriately removing both TAGs for the same YKO is less than 0.000 01. Using this procedure we defined 193 DNTAGs (purple circles) and 244 UPTAGs (blue circles) as primer-batch specific outliers. Six YKOs had both UP and DN ratios filtered (orange circles).

Next, we considered the effect of TAG-specific hybridization behavior resulting from the presence of nucleotide mutations found in some of the TAGs and universal

priming sites (Eason et al. 2004). This is important because sensitivity will be markedly

affected when TAGs fail to provide a meaningful measure of strain representation.

Histograms of $\log_2$ signal intensity display a bimodal distribution (Figure 2b and data not

shown) for UP- and DNTAGs whether Cy3 or Cy5 labeling is used. The lower peak is

close to background intensities and contains nonfunctional TAGs with absent or

inefficient hybridization. While TAG sequence discrepancies have been characterized

(Eason et al. 2004), knowledge of the presence and nature of mutations was insufficient

to fully predict hybridization behavior (Yuan et al. 2005; Eason et al. 2004).

The naïve approach to summarizing UP- and DNTAG information is to average

their observed log ratios. This solution will yield suboptimal measures when one of the

TAGs is non-functional. We propose a procedure exploiting the bimodal distribution of

TAG intensities to improve on simple averaging. To determine if a TAG is non-

functional we fit a mixture model, as in Irizarry et al. (2003), to the log intensity data for

the control sample. The model fits two normal distributions to the Cy5 data, one for the

lower mode and one for the upper mode. The 'blank' (YQL) features (Yuan et al. 2005)

define the location and width (mean and SD) of the lower distribution. With this fitted

model in place, we can predict the probability that each TAG is 'present' (Figure 2b). We

consider a DN/UPTAG non-functional when it is predicted absent while the

complementary UP/DNTAG is present. We define a weighted average $= w * \mathrm{UP} + (1 - w)$

$* \mathrm{DN}$, where $w = 0.5 + [P(\mathrm{UP\ present}) - P(\mathrm{DN\ present})]/2$. Thus, when UP is present

($P_{\mathrm{UP}} = 1$) and DN is absent ($P_{\mathrm{DN}} = 0$), $w = 1$ and only UPTAG is used (Figure 2c). We

describe a less complex procedure in Supplementary Note 1 that uses binary absent or

present values ($P = 0$ or 1) and performs similarly (data not shown). Researchers using unsophisticated analysis software such as spreadsheet applications may prefer the simple procedure.

We compared the performance of these two strategies and use of UP- or DNTAGs alone with Receiver Operating Characteristic (ROC) curves based on the spike-in experiment. For this analysis, nominal ratios below 2-fold were excluded from the list of True Positives. This choice is appropriate because 2-fold representation difference corresponds to a subtle growth defect at the margin of detection in colony measurement (1.25-fold colony diameter difference is predicted by hemispherical colony volume = $2\pi r^3/3$). Supplementary Figures 2–3 present ROC curves with varying stringencies for inclusion as 'True Positive', including every spiked-in YKO (1.26-fold or higher). ROC curves in the range of false positives likely to be acceptable (Figure 3a and Supplementary Figure 2) demonstrate that the artifact filtering process has a significant effect on specificity (Supplementary Figure 4 shows the full ROC curves). Additional filtering of non-functional TAGs by the weighted average improves results further (Figure 3b and Supplementary Figure 3).

The effect of two filters, one which removes the systematic artifacts and a second which removes non-hybridizing TAGs, is demonstrated by ratio-intensity plots. A naïve approach to analysis would average UP and DN log ratio to produce a measure M for relative strain representation. By filtering systematic artifacts, noise is significantly reduced (Figure 3c and d, open circles and Supplementary Figure 5). Additionally, combining UP and DN selectively provides increased sensitivity for a number of spiked-

in strains (filled shapes).

Discussion

In summary, we present a spike-in pool design that allows evaluation of various

methods for generating measures of differential strain representation. Using this

experiment, we determined that the largest factor affecting specificity is the presence of

primer batch-specific artifacts, evident in control self-self hybridizations. These artifacts

may result from extremely low levels of contaminating TAG sequence template

introduced before the labeling PCR. Accidental introduction of contaminants may occur

at multiple steps, including the high-performance liquid chromatography (HPLC) column

purification of Cy5- and Cy3-labeled primers at their manufacture as well as laboratory

manipulation of primer batches during initial stock and aliquot preparation. Because the

artifacts are consistent only within batches of primer sets, contamination must occur at

initial preparation or during manufacture. Yuan et al. (2005) discuss the unusually large

dilutions required to prevent contamination in TAG labeling reactions. While the source

of these artifacts is uncertain, there are several options for minimizing their effect. The

approach we present uses a control hybridization of one DNA sample labeled with both

primer sets, such that every TAG is present at equal amounts in the two labeling

reactions. Deviations from expected 1:1 ratio can be recognized and filtering is applied.

The methods we describe improve detection of true signal difference between

samples; however, they are not perfect. Once primer-batch specific artifacts are removed,

noise is increased slightly with the weighted average method compared to averaging (see

Supplementary Figure 5d and e). Additionally, the weighted average could cause

decreased sensitivity when cross-hybridization occurs for one TAG from a low represented strain. The TAG that accurately reflects the low representation of the YKO may be discounted while the cross-hybridizing TAG is emphasized. These problems could be minimized by improving the criteria for selecting a TAG as non-hybridizing, perhaps by examining behavior across many microarrays. The advantage of this method is that it requires as few as two microarray hybridizations (self-self and experiment) to perform well. We have tested these methods to provide optimal results from SLAM experiments, where YKOs that decrease in representation from control to experimental samples are sought. However, appropriately applied, other TAG microarray experiments should benefit from the procedures we describe.

Acknowledgments

References

Arevalo-Rodriguez M, Pan X, Boeke JD, Heitman J. 2004. FKBP12 controls aspartate pathway flux in *Saccharomyces cerevisiae* to prevent toxic intermediate accumulation. *Eukaryot Cell* 3(5):1287–96.

Dudoit, S, Yang, YH, Callow, MJ, Speed, TP. 2002. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments.

Statistica Sinica 12(1):111–39.

Eason RG, Pourmand N, Tongprasit W, Herman ZS, Anthony K, Jejelowo O, Davis RW, Stolc V. 2004. Characterization of synthetic DNA bar codes in *Saccharomyces cerevisiae* gene-deletion strains. *Proc Natl Acad Sci USA* 101(30):11046–51.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418(6896):387–91.

Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 18 Suppl 1:S96–104.

Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. *J Comput Graph Stat* 5(3):299–314.

Irizarry RA, Ooi SL, Wu Z, Boeke JD. 2003. Use of mixture models in a microarray-based screening procedure for detecting differentially represented yeast mutants. *Stat Appl Genet Mol Biol* 2:Article1.

Kerr MK, Martin M, Churchill GA. 2000. Analysis of variance for gene expression microarray data. *J Comput Biol* 7(6):819–37.

Kooperberg C, Fazzio TG, Delrow JJ, Tsukiyama T. 2002. Improved background correction for spotted DNA microarrays. *J Comput Biol* 9(1):55–66.

Lee MS, Spencer FA 2004. Bipolar orientation of chromosomes in *Saccharomyces cerevisiae* is monitored by Mad1 and Mad2, but not by Mad3. *Proc Natl Acad Sci USA*

101(29):10655–60.

Ooi SL, Shoemaker DD, Boeke JD. 2001. A DNA microarray-based genetic screen for nonhomologous end-joining mutants in *Saccharomyces cerevisiae*. *Science* 294(5551):2552–6.

Ooi SL, Shoemaker DD, Boeke JD. 2003. DNA helicase interaction network defined using synthetic lethality analyzed by microarray. *Nat Genet* 35(3):277–86.

Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD. 2004. A robust toolkit for functional profiling of the yeast genome. *Mol Cell* 16(3):487–96.

Shoemaker DD, Lashkari DA, Morris D, Mittmann M, Davis RW. 1996. Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* 14(4):450–6.

Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, et al. 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294(5550):2364–8.

Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, et al. 2004. Global mapping of the yeast genetic interaction network. *Science* 303(5659):808–13.

Warren CD, Eckley DM, Lee MS, Hanna JS, Hughes A, Peyser B, Jie C, Irizarry R, Spencer FA. 2004. S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion. *Mol Biol Cell* 15(4):1724–35.

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, et al. 1999. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901–6.

Yang YH, Buckley MJ, Dudoit S, Speed TP. 2002. Comparison of methods for image analysis on cDNA microarray data. *J Comput Graph Stat* 11(1):108–36.

Yuan DS, Pan X, Ooi SL, Peyser BD, Spencer FA, Irizarry RA, Boeke JD. 2005. Improved microarray methods for profiling the Yeast Knockout strain collection. *Nucleic Acids Res* 33(12):e103.
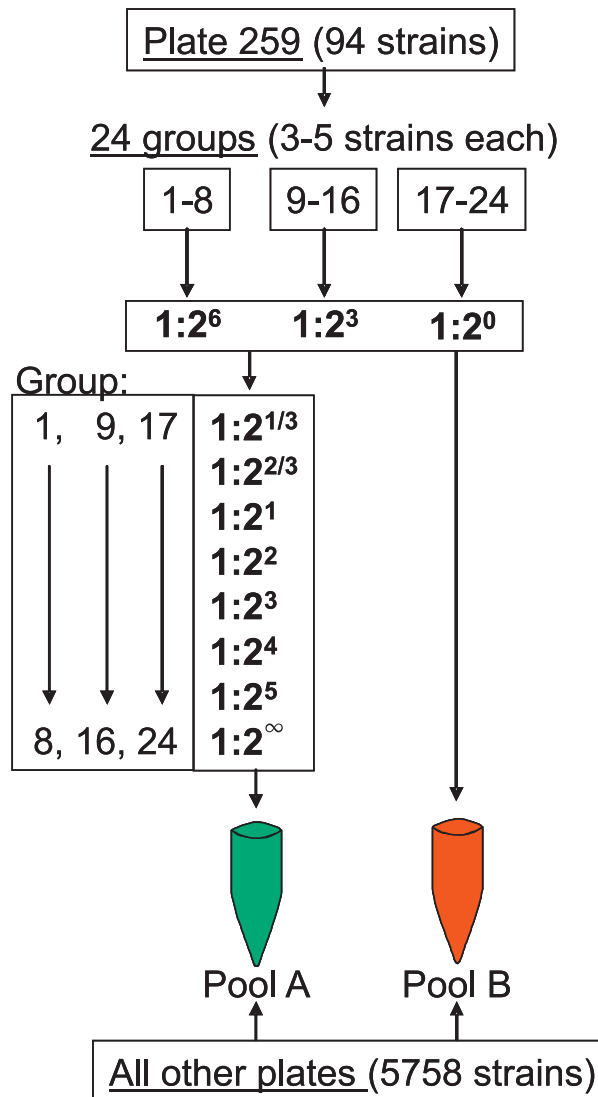
Figure 1. Design of spike-in pools. Two pools were created ('A' and 'B') such that 94 strains were differentially represented between the two pools. The 94 differentially represented strains were diluted 1:1, 1:8 or 1:64 (High, Medium or Low representation), then added to Pool B. Each strain was then diluted again from $1:2^{1/3}$ to $1:2^5$ and added to Pool A. One set of strains from each representation group was not added to Pool A (dilution $1:2^\infty$).

Figure 2. Development of TAG filters. (a) Self-self hybridizations. Pool A gDNA

was used as template for TAG labeling reactions with each primer set (UP/DNTAG

Cy5/Cy3). Median values across three experiments were displayed. Each point represents

a single YKO. Red dotted lines are three SDs. Blue circles, artifacts specific to UP ratio;

purple circles, artifacts specific to DN ratio; orange circles, artifacts in both TAGs. (b)

Histogram of $\log_2$ UPTAG Cy5 signal values from pool A versus B hybridization.

Results from DNTAG and Cy3 are similar. Red line, probability each feature belongs to

the upper (righthand) distribution. (c) UPTAG Cy5 versus DNTAG Cy5. Each point

represents a single YKO. Therefore, for each point the numbers of UP- and DNTAGs in

the sample are identical. Red lines, values at which absent/present probabilities equal 0.5.

In the weighted average method, DN ratio was weighted higher for YKOs in the upper

left quadrant, and UP ratio was weighted higher for YKOs in the lower right quadrant.
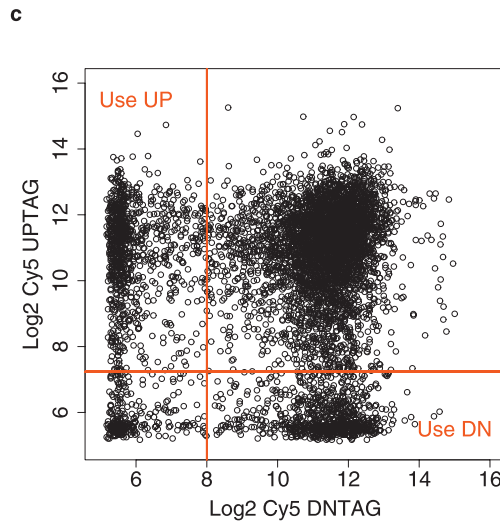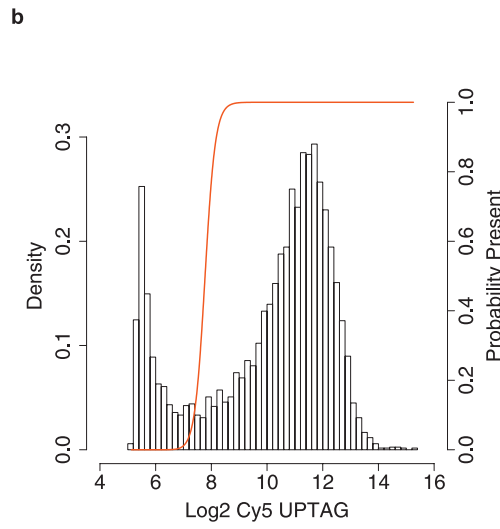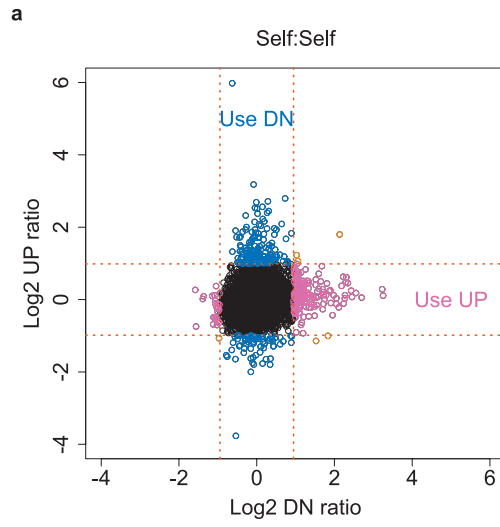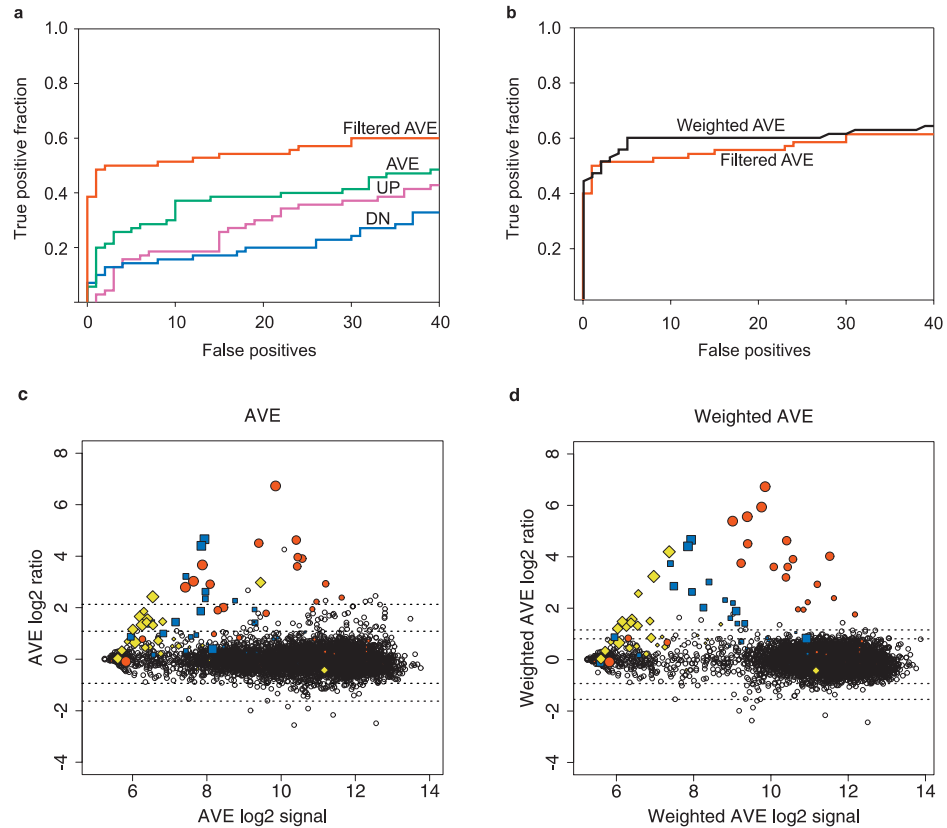
Figure 3. Measures of sensitivity/specificity. (a and b) ROC for several methods of calculating relative YKO representation. True Positive is defined as any YKO with known pool B:A ratio > 2. False Positive is defined as any YKO with known B:A ratio of 1. ROC curves for UP ratio alone (purple line), DN ratio (blue line), Average ratio (green line), Filtered average ratio (red line) and weighted average ratio (black line). (c and d) Ratio-intensity plots using simple averaging of UP and DN ratios or weighted average ratios. Black open circles, YKOs with known B:A ratio of 1; filled circles, high representation YKOs; squares, medium representation YKOs; diamonds, low representation YKOs. Point size is related to known B:A ratio: from small to large, $2^{1/3}$, $2^{2/3}$, 2, 2^2, 2^3, 2^4, 2^5 and 2^∞. Black dotted lines are 0.1, 1, 99 and 99.9 percentiles for YKOs with known B:A ratio = 1.

38

Supplementary note 1

An alternate weighting procedure can be used where p equals 0 or 1, such that all control (Cy5) TAGs above a threshold are considered 'present' (p=1) while TAGs below the threshold are 'absent.' The threshold value for each array can be calculated using the mean plus three standard deviations for $\log_2$ Cy5 intensity values of expected 'blank' features. On the "Hopkins TAG Array" (GEO accession GPL1444) the 'YQL' features can be used for this purpose. On other TAG arrays, researchers could use features representing essential yeast knockout strains (YKOs), if the sample is from a haploid pool. Once the threshold value for a given array is determined, each UPTAG and DNTAG is annotated 'present' (p=1) or 'absent' (p=0) based on the control channel intensity, and the corresponding log ratios are averaged for each YKO using $w * UP +$ $(1 - w) * DN$, where $w = 0.5 + (p(UP\ present) - p(DN\ present))/2$.
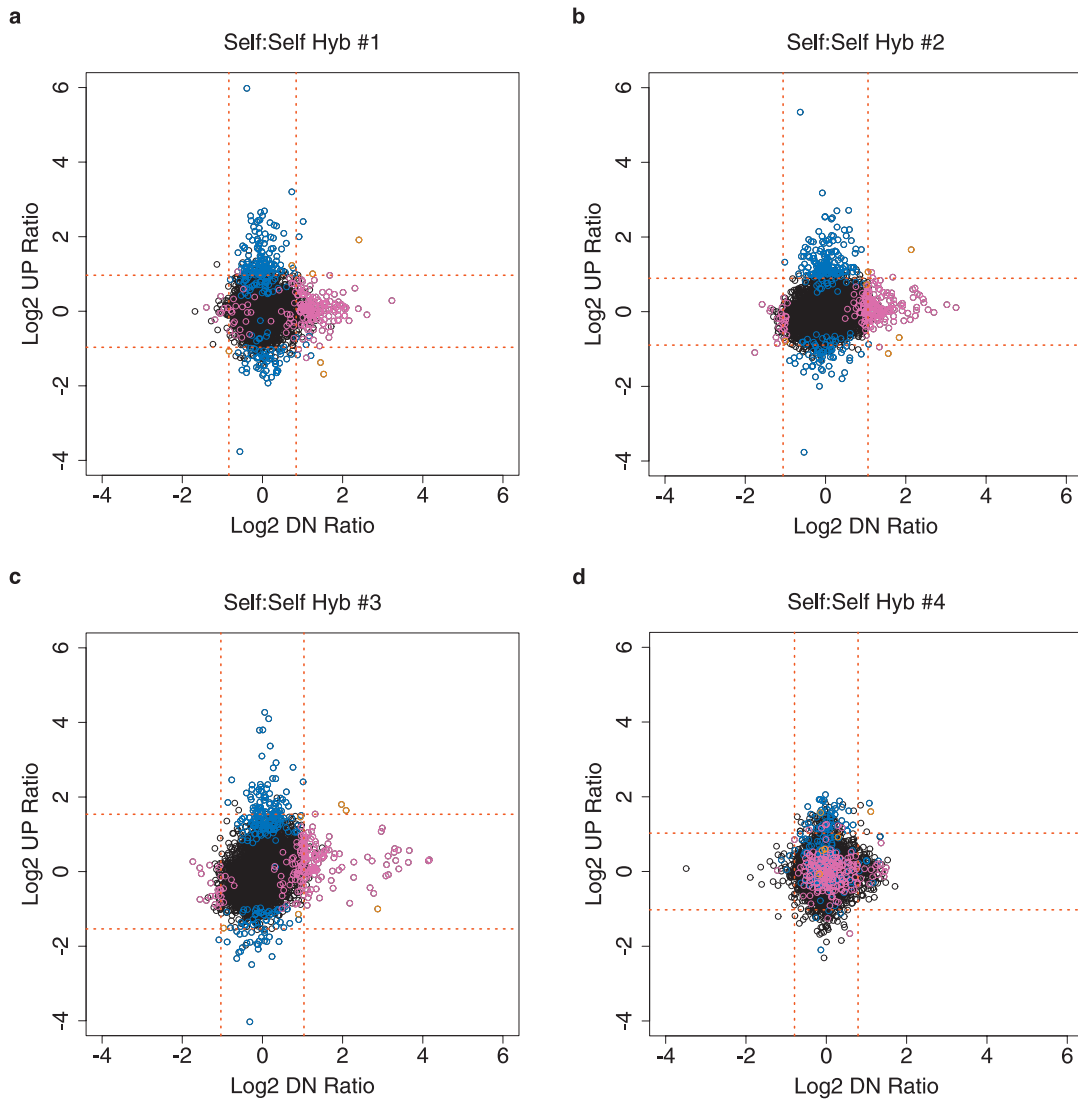
Supplementary Figure 1. UPTAG $\log_2$ ratio versus DNTAG $\log_2$ ratio. Three independent self:self hybridizations were performed using the same batch of Cy3/Cy5 labeled primers (a-c). A different batch of primers was used to perform the hybridization in panel d. Each point represents a single yeast knockout strain. Blue circles, UPTAG specific artifacts from Figure 1; purple circles, DNTAG specific artifacts; orange circles, artifacts in both tags. Red lines, three times S.D. Identities of colored points are constant across all panels.
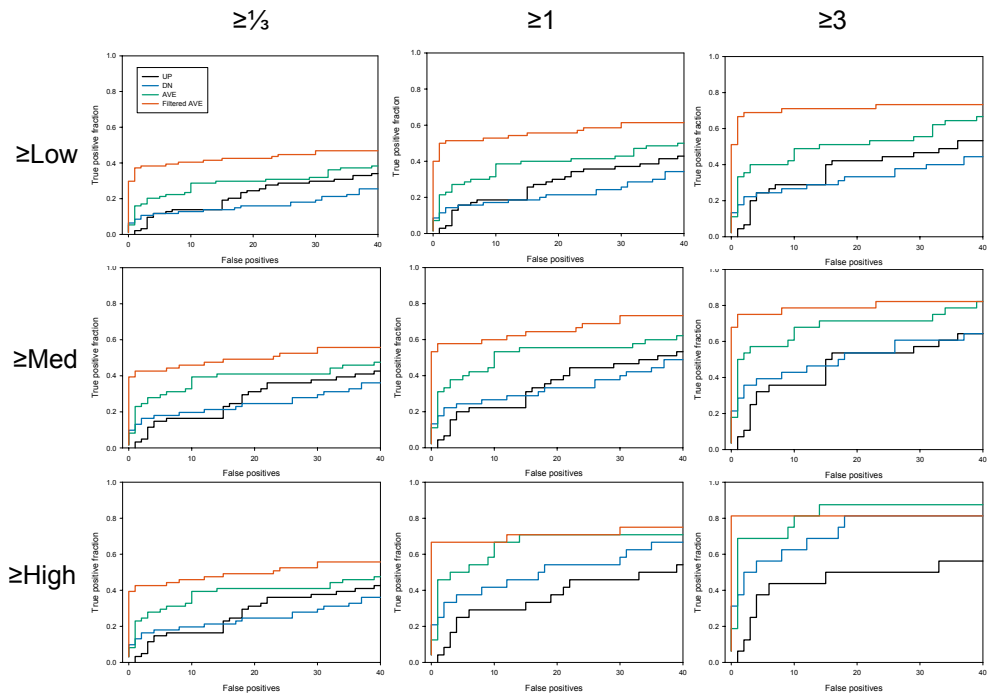
Supplementary Figure 2. ROC curves. True positives defined as all knockout strains with $\log_2$ pool B:A ratio of ⅓ and greater (left column), 1 and greater, (middle column), or 3 and greater (right column), and pool B representation level of 1:64 or greater (top row), 1:8 or greater (middle row) or 1:1 (bottom row). False positives defined as all knockout strains present at 1:1 ratio. UP ratio, black lines; DN ratio, blue lines; average ratio, green lines; filtered average ratio, red lines.
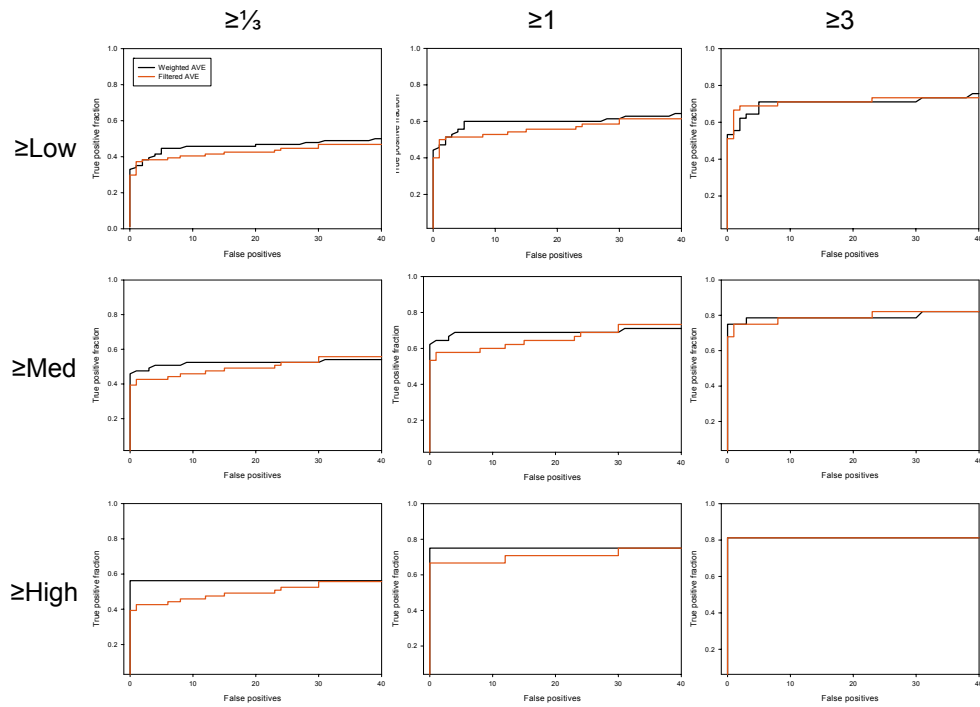
Supplementary Figure 3. ROC curves. True positives defined as all knockout

strains with log$_2$ pool B:A ratio of ⅓ and greater (left column), 1 and greater, (middle

column), or 3 and greater (right column), and pool B representation level of 1:64 or

greater (top row), 1:8 or greater (middle row) or 1:1 (bottom row). False positives defined

as all knockout strains present at 1:1 ratio. Weighted average ratio, black lines; filtered

average ratio, red lines.
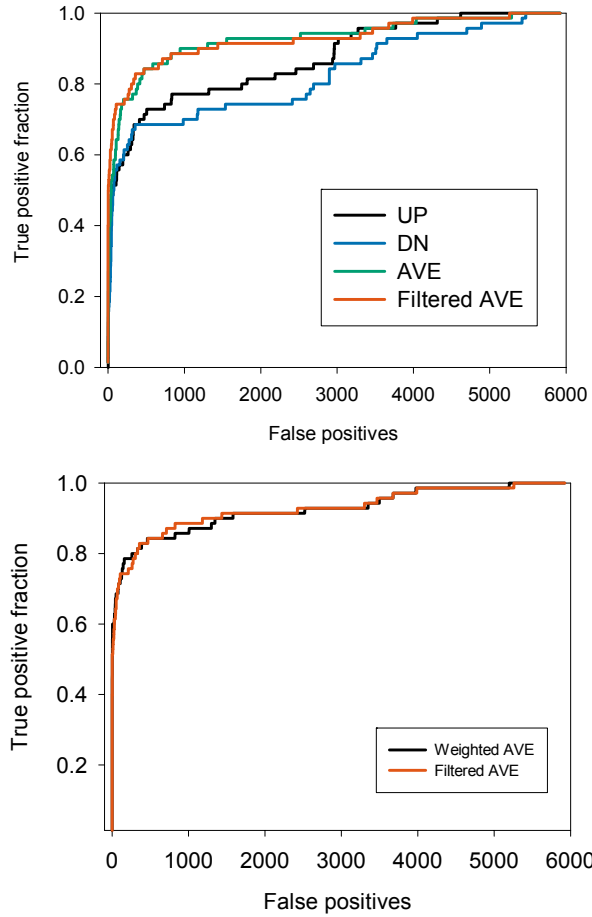
Supplementary Figure 4. ROC curves for several methods of calculating relative YKO representation. True Positive is defined as all knockout strains with $\log_2$ pool B:A ratio ≥ 1. False Positive is defined as all knockout strains present at 1:1 ratio. (a) ROC curves for UP ratio alone (black line), DN ratio (blue line), Average ratio (green line), and Filtered AVE ratio (red line). (b) Weighted AVE ratio (black line), and Filtered AVE ratio (red line).
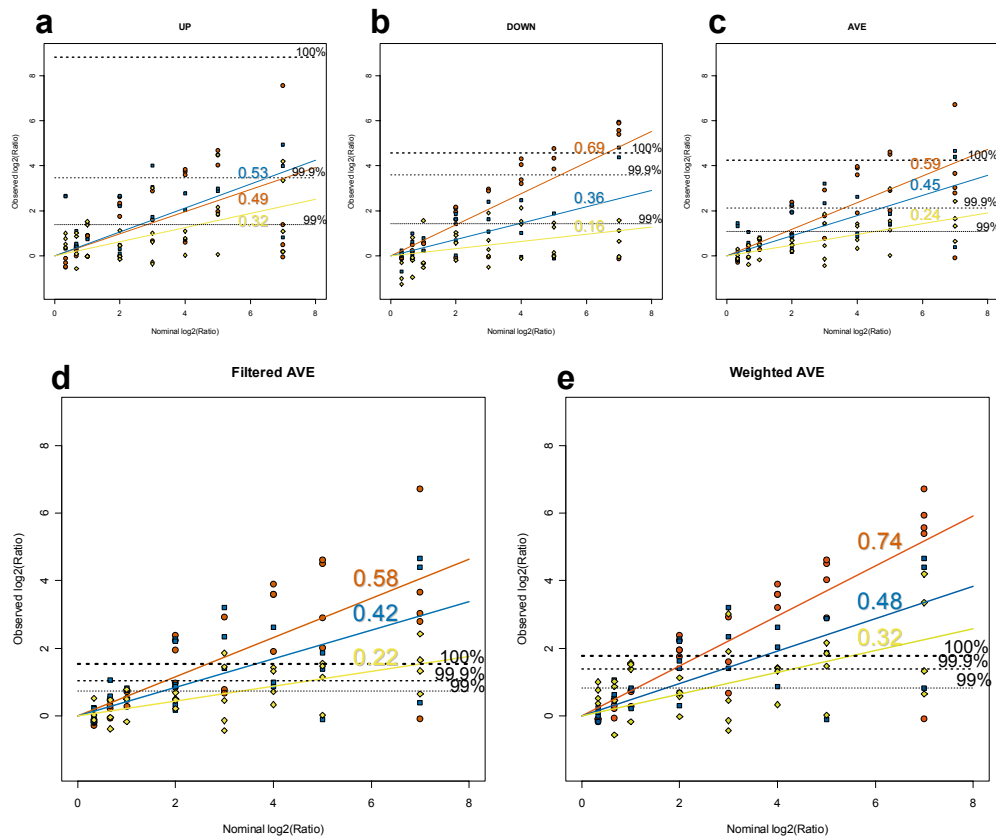
Supplementary Figure 5. Detection slopes. Five methods for generating a single log ratio measure of differential pool B/A representation were analyzed by comparing observed ratio to known ratio. Points from High (red points and lines), Medium (blue), and Low (yellow) representation groups were used to generate three linear fits (least squares). Slope of the fitted line is indicated in corresponding color adjacent each line. Dotted lines indicate the noise range of all ratios for known 1:1 knockout strains. Panels a-e: UP ratio alone, DN alone, AVE ratio, AVE ratio with artifacts removed, Weighted AVE ratio with artifacts removed.

Supplementary Table 1. Representation levels of heterozygous deletion strains incorporated into the spike-in pools A and B.

ORF	Strain Plate	Strain Row	Strain Column	log$_2$(Relative Conc. in B)	log$_2$(Relative Conc. in A)	log$_2$(B/A)
YDL064W	259	A	1	0	-0.33	0.33
YDL065C	259	A	2	0	-0.33	0.33
YDL066W	259	A	3	0	-0.33	0.33
YDL067C	259	A	4	0	-0.33	0.33
YDL068W	259	A	5	0	-0.67	0.67
YDL069C	259	A	6	0	-0.67	0.67
YDL070W	259	A	7	0	-0.67	0.67
YDL071C	259	A	8	0	-0.67	0.67
YDL072C	259	A	9	0	-1	1
YDL073W	259	A	10	0	-1	1
YDL074C	259	A	11	0	-1	1
YDL075W	259	A	12	0	-2	2
YDL076C	259	B	1	0	-3	3
YDL077C	259	B	2	0	-3	3
YDL078C	259	B	3	0	-3	3
YDL079C	259	B	4	0	-4	4
YDL080C	259	B	5	0	-5	5
YDL081C	259	B	6	0	-inf	inf
YDL082W	259	B	7	-3	-3.33	0.33
YDL083C	259	B	8	-3	-3.33	0.33
YDL084W	259	B	9	-3	-3.33	0.33
YDL085W	259	B	10	-3	-3.33	0.33
YDL086W	259	B	11	-3	-3.67	0.67
YDL087C	259	B	12	0	-2	2
YDL088C	259	C	1	-3	-4	1
YDL089W	259	C	2	-3	-4	1
YDL090C	259	C	3	-3	-4	1
YDL091C	259	C	4	0	-4	4
YDL092W	259	C	5	0	-5	5
YDL093W	259	C	6	0	-inf	inf
YDL094C	259	C	7	-3	-5	2
YDL095W	259	C	8	-3	-6	3
YDR437W	259	C	9	-3	-5	2
YDR438W	259	C	10	-3	-7	4
YDR439W	259	C	11	-3	-3.67	0.67
YDR440W	259	C	12	0	-2	2
YDR441C	259	D	1	-3	-8	5

ORF	Strain Plate	Strain Row	Strain Column	log$_2$(Relative Conc. in B)	log$_2$(Relative Conc. in A)	log$_2$(B/A)
YDR442W	259	D	2	-3	-8	5
YDR443C	259	D	3	-3	-8	5
YDR446W	259	D	4	0	-4	4
YDR447C	259	D	5	0	-5	5
YDR448W	259	D	6	0	-inf	inf
YDR449C	259	D	7	-3	-5	2
YDR450W	259	D	8	-3	-6	3
YDR451C	259	D	9	-3	-6	3
YDR452W	259	D	10	-3	-7	4
YDR453C	259	D	11	-3	-3.67	0.67
YDR454C	259	D	12	0	-2	2
YDR455C	259	E	1	-3	-inf	inf
YDR456W	259	E	2	-3	-inf	inf
YDR457W	259	E	3	-3	-inf	inf
YDR458C	259	E	4	0	-4	4
YDR459C	259	E	5	0	-5	5
YDR460W	259	E	6	0	-inf	inf
YDR462W	259	E	7	-3	-5	2
YDR463W	259	E	8	-3	-5	2
YDR464W	259	E	9	-3	-5	2
YDR465C	259	E	10	-3	-7	4
YDR466W	259	E	12	0	-2	2
YDR467C	259	F	1	-6	-6.33	0.33
YDR468C	259	F	2	-6	-6.67	0.67
YDR469W	259	F	3	-6	-6.67	0.67
YDR470C	259	F	4	-6	-7	1
YDR471W	259	F	5	-6	-7	1
YDR472W	259	F	6	0	-inf	inf
YDR473C	259	F	7	-6	-8	2
YDR474C	259	F	8	-6	-8	2
YDR475C	259	F	9	-6	-8	2
YDR476C	259	F	10	-6	-8	2
YDR477W	259	F	11	-3	-3.67	0.67
YDR478W	259	F	12	-6	-9	3
YDR479C	259	G	1	-6	-6.33	0.33
YDR480W	259	G	2	-6	-6.67	0.67
YDR481C	259	G	3	-6	-6.67	0.67
YDR482C	259	G	4	-6	-7	1
YDR483W	259	G	5	-6	-7	1
YDR484W	259	G	6	-6	-10	4

ORF	Strain Plate	Strain Row	Strain Column	log$_2$(Relative Conc. in B)	log$_2$(Relative Conc. in A)	log$_2$(B/A)
YDR485C	259	G	7	-6	-10	4
YDR486C	259	G	8	-6	-10	4
YDR487C	259	G	9	-6	-11	5
YDR488C	259	G	10	-6	-11	5
YDR489W	259	G	11	-6	-9	3
YDR490C	259	G	12	-6	-9	3
YDR491C	259	H	1	-6	-6.33	0.33
YDR492W	259	H	2	-6	-6.33	0.33
YDR494W	259	H	3	-6	-inf	inf
YDR495C	259	H	5	-6	-inf	inf
YDR496C	259	H	6	-6	-inf	inf
YDR497C	259	H	7	-6	-inf	inf
YDR499W	259	H	8	-6	-10	4
YDR500C	259	H	9	-6	-11	5
YDR503C	259	H	10	-6	-11	5
YDR504C	259	H	11	-6	-9	3
YDR505C	259	H	12	-6	-9	3
All other strains				0	0	0

Chapter 4.

Synthetic lethal interactions predicted by analysis of high-throughput

dSLAM.

Abstract

Using analysis of microarray data from 707 unique dSLAM query genes, we predict 1690 synthetic lethal interactions between gene pairs in *S. cerevisiae*, including 1410 previously unknown. This analysis makes use of data from double mutants generated in two orientations—*yko1Δ::URA3* (query) *yko2Δ::KanMX* (target) and *yko2Δ::URA3* (query) *yko1Δ::KanMX* (target)—to improve the predictive ability of microarrays. We compare several methods for using bidirectional information with known interactions available from BioGRID to estimate sensitivity and specificity. Currently, we present data from 707 query genes, but all nonessential yeast knockout mutants are planned. The value of this bidirectional approach should increase as more data become available.

Introduction

Synthetic lethality provides insight into the mechanisms of robustness found in living systems (Hartman et al. 2001; Tucker and Fields 2003; Wagner 2005). Synthetic lethal pairs are two alleles that are individually nonlethal but cause lethality when combined. Also considered are interactions where a growth defect more severe than expected is caused by interaction of two alleles. This has been called "synthetic fitness defect" or "synthetic sick," though in some cases "synthetic semi-lethality" may be apropos after Dobzhansky (1946). Interaction between two point mutant alleles often coincides with physical interactions between proteins encoded by those genes. In contrast, synthetic interactions found between null alleles rarely coincide with physical interaction between the two genes (Tong et al. 2004; Kelley and Ideker 2005; Ye and

49

Peyser et al. 2005). These instead represent functions that are required in absence of a compensating pathway.

With introduction of the yeast knockout (YKO) collection (Winzeler and Shoemaker et al. 1999), it became possible to readily probe the near-complete yeast genome for synthetic interactions between null alleles. Several approaches have been applied to the task, including synthetic genetic array (SGA), epistatic miniarray profile (E-MAP), synthetic lethality analyzed by microarray (SLAM), and diploid-based SLAM (dSLAM) (Tong et al. 2001; Schuldiner et al. 2005; Ooi et al. 2003; Pan et al. 2004). In SLAM experiments, the two molecular barcodes (TAGs) incorporated into each deletion (Shoemaker et al. 1996), called UPTAG and DNTAG, are simultaneously interrogated to estimate changes in strain abundance between conditions. All UPTAGs or all DNTAGs can be amplified from a sample using universal flanking primer sites.

In dSLAM, pooled heterozygous diploid YKOs are transformed with a second "query" null mutation. These diploids containing query and target null alleles in heterozygous condition are then sporulated, and haploid *MAT*a cells are selected with or without requirement for the query allele (experiment and control, respectively). The cells that survive selection are then processed for genomic DNA and used to prepare labeled TAGs. We present results from high-throughput dSLAM experiments, and predict new interactions based solely on these microarray data.

Methods

Synthetic lethality screens were performed generally as described by Pan et al.

(2007). Briefly, pooled haploid-convertible YKO mutants were transformed *en masse* with a deletion construct consisting of ~1.5 kb up- and downstream sequence for the query ORF surrounding a URA3 cassette. Transformants were selected on −uracil media, scraped and sporulated. *MAT*a haploids were selected on "magic" haploid selection media +G418 (control) and −uracil +G418 (experiment).

Genomic DNA purified from each selection was used as template for TAG labeling reactions as described (Yuan et al. 2005). TAGs from control samples were labeled with Cy5, and experiment samples were labeled with Cy3. Labeled TAGs were hybridized to custom microarrays ("Hopkins Tag Array" from Agilent, Gene Expression Omnibus accession number GPL1444) as previously described, and scanned using a GenePix scanner (Axon Instruments).

Microarray results were stored as GenePix Results (".gpr") files and organized into subsets by primer batch. Each set of scans from a single primer batch was analyzed using R, an open-source data processing environment (Ihaka and Gentleman 1996). The *limma* package (Smyth 2005) of BioConductor (http://www.bioconductor.org) was used for data structures, normalization, and generation of moderated *t*-statistics for replicate data using the empirical Bayes procedure (Smyth and Speed 2003). The array data were obtained from the GenePix median pixel intensity values. Arrays were normalized and background subtracted using the "loess" and "normexp" methods of the "normalizeWithinArrays()" function in *limma*, respectively (Smyth 2004). The background subtraction was performed with an offset of 16 to reduce variance explosion at low intensities. Array average intensities were then normalized using the "Aquantile" method of the

51

"normalizeAcrossArrays()" function. Data for UPTAG and DNTAG were treated separately in all cases. Deletions of ORFs adjacent to the query can cause artificially high control/experiment ratios because of interference with the targeting of the deletion construct to the chromosome homologue carrying the knockout, and subsequent repulsion during meiosis due to linkage. Therefore, ORFs within 2 of either side of the query were assigned a Z-score of 0. Once Z-scores were produced for each array, the results of all primer sets were combined and analysis was continued using R.

Results

We performed dSLAM experiments on 707 unique query open reading frames (ORFs) using procedures as described in Pan et al. (2007). DNA from double mutant cells was used to label TAGs with Cy3, while Cy5 was used to label TAGs from control single-mutant DNA. Here, large values for control/experiment (C/E) ratio are expected for strains that do not survive deletion of both ORFs. In general, normalization procedures were similar to those previously described (Peyser et al. 2005). See Methods for a complete description of normalization and background correction. As shown in Peyser et al. (2005), two problems associated with TAG arrays are primer-batch–specific artifacts, and poorly hybridizing TAGs. In this work, we apply new techniques for identifying bad TAGs and primer artifacts using many microarray hybridizations. In addition, we make use of information from two knockout orientations—*yko1Δ::URA3* (query) *yko2Δ::KanMX* (target) and *yko2Δ::URA3* (query) *yko1Δ::KanMX* (target)—to increase specificity for each gene pair.

Broken TAGs

A priori, signal from TAGs that provide no information should never change, and variation seen for these TAGs should be solely due to noise. We examined a large number of TAG hybridizations to define the variability of each TAG and remove those with extremely small standard deviation (SD). Data from 1121 scans were quantile normalized (Bolstad et al. 2003) across all scans and both colors (Cy5 and Cy3) for UPTAG and DNTAG separately, without regard to primer batch. Following normalization, the average $\log_2$ signal intensity was plotted with a robust estimate of the SD (median absolute deviation, MAD) of the $\log_2$ intensities (Figure 1). As expected, extremely small values for MAD are associated with low signal intensity. We applied a cutoff to the MAD values, at approximately the midpoint between two modes in the distributions (0.40 for UPTAG and 0.35 for DNTAG), and annotated all TAGs below these cutoffs as "failed" (see Supplementary Table 1). If these bad TAGs provide no information about the target molecules, their data will consist of only noise. The identification of failed TAGs permits removal of these data to reduce this noise. Additionally, some YKOs (192 strains) were created with no DNTAG. For convenience, these strains were also annotated to have bad DNTAGs.

Figure 2 displays density histograms of TAG behavior by type across multiple microarrays. The distribution of bad TAGs coincides with negative control features included on the arrays (see Yuan et al. 2005). Additionally, while essential mutants should not grow on haploid selective media, the TAG signals for those mutants are detectable on microarrays, presumably due to presence of dead or dying cells. This distribution is bimodal before removal of bad TAGs, which do not provide information

about the low levels of essential mutants. Note that bad TAGs as defined in Figure 1 simply reproduce the negative control distribution.

Removing artifacts

Most importantly, primer-batch–specific artifacts cause spurious results for some TAGs. These artifacts are generally consistent within a single batch of labeled primer, but vary between batches (Peyser et al. 2005). One remedy is to remove the TAGs with artifactual signal from analysis. In contrast, here we apply a transformation to the data that expresses changes in TAG levels between experimental conditions as change from typical TAG $\log_2$ ratios within each primer batch. This assumes that YKOs usually do not change abundance between experimental conditions, and that $\log_2$ ratios that are not 0 on average are influenced by artifacts. We know this is false for some strains that respond to the uracil selection without regard to the query deletion (see Table 1). However, since this uracil effect is not the desired biological phenotype, its removal is also beneficial. We define typical TAG behavior by using the mean and SD of the $\log_2$ C/E ratio for each TAG among a set of many hybridizations performed with the same primer batch. Again, we use the MAD as a robust estimate for SD. The data are transformed to a Z-score, which is the number of SDs from the mean. Figure 3 shows the procedure for two TAGs, one of which is typical, and one of which displays a primer-batch–specific artifact. The Z-score procedure re-centers the distributions and equalizes the variance. This method is successful at equalizing variance across intensities, as well (see Durbin and Rocke 2002; Huber et al. 2002).

The Z-score procedure improves on the filtering method presented in Peyser et al.

54

(2005) by retaining data that would otherwise be removed. For example, the TAG represented by the red line in Figure 3 displays $\log_2$ ratio higher than its typical behavior on 1 microarray, corresponding to a Z-score of about 7 (arrow). This TAG would have been removed from analysis with the previously described filtering procedure since the average $\log_2$ ratio is much greater than 0. With this Z-score method, the potentially useful information is retained.

Combining UP- and DNTAG information

Each knockout strain (except for 192 strains lacking DNTAGs) has information from two TAGs. However, when a TAG feature provides no information about the abundance of target molecules, including the corresponding data will provide only noise. Therefore, when combining UP- and DNTAG information, only functional TAGs should be retained.

When combining independent normal distributions with mean and SD of 0 and 1, adding the values then dividing by the square root of 2 results in a distribution that is also mean 0 and SD of 1. In Figure 4, this can be understood as a measure of the distance along the red line from the origin to a perpendicular line that intersects the point. When TAGs are combined in this manner, a standard normal distribution can be maintained by substituting the UP- or DNTAG value for the combined value when any corresponding DN- or UPTAG is annotated "failed." This method weights strains with agreement for UPTAG and DNTAG higher than strains with only one functional TAG, but less than the double weight provided by adding Z-scores.

55

Combining information from two arrays

Significant noise is detectable in microarray experiments when comparing biological replicates. Figure 5 shows Z-scores from two of three independent dSLAM screens performed with the same query. While the same experiment was performed, the results are not highly correlated ($r = 0.15$). However, low p-values (warm colors) produced using the empirical Bayes moderated t-statistic (Smyth 2004) are associated with results where two arrays agree and are enriched for true positive interactions (Tong et al. 2004), indicated by black filled circles. This level of variation in microarray experiment results is not unique to TAG microarrays, and poses a problem for the high-throughput project since only one experiment is planned for each query ORF.

One feature of $yko1\Delta$ $yko2\Delta$ interactions is that interesting results should occur both where $yko1\Delta$ serves as query and when $yko2\Delta$ does. Here we investigate the utility of information for deletions made in both orientations. While we were unable to produce p-values as indicated in Figure 5 with only two experiments, we could use the additional information for each interaction from the corresponding gene pair.

One method to combine the data from two orientations would be to average the values in the same manner as UP- and DNTAG were combined. Another possibility is to keep UP- and DNTAG values separate and choose the median of up to four values. This would reduce the impact of a single spuriously extreme value. Notice that both of these methods both make it possible to predict interactions for gene pairs where one mutant strain has no working TAGs.

We present a method that views agreement in the Z-scores from the two orientations

56

as confirmatory, and disagreement as evidence of noise. This is shown in Figure 6, which displays the combined UP- and DNTAG Z-score for each orientation of a given gene pair. As expected, points are more dense in the upper right quadrant, consistent with a real synthetic lethal effect detected in two microarrays. If points in the lower right quadrant are a good representation of noise, the two quadrants can be compared to estimate the false discovery rate (FDR). We perform this estimation by creating hyperbolic cutoffs as shown in Figure 7, and counting the number of points present above corresponding cutoffs in each quadrant. Thus, when a cutoff generates 100 predicted SL pairs in the upper right quadrant, and the corresponding cutoff generates 25 expected false interactions in the lower right quadrant, the estimated FDR is 0.25 at that cutoff. If we knew the true phenotypes of all 100 pairs above that cutoff, we would expect that 75 are true, and 25 are above the cutoff simply due to noise. This estimated FDR should be conservative, since not all points in the lower right quadrant will be truly false.

We can similarly perform this procedure for gene pairs in the lower left quadrant. Here, two gene deletions together appear to improve growth compared to the individual deletions. This phenotype is termed Synthetic Rescue, though it historically applied to a lethal allele becoming nonlethal in presence of another allele. We apply it to nonessential gene deletions that display improved fitness in response to deletion of a second gene.

One weakness of this method is that YKO strains with no functional TAGs cannot yield predicted interactions involving that strain. This is the case even when the ORF is used as the query. Thankfully only 78 nonessential heterozygous diploid strains, 7 of which are included in this analysis, have both TAGs flagged as bad. For these strains, the

Z-scores can be used directly to generate candidate lists, and additional experiments can be performed to improve predictions.

FDR method outperforms alternatives

For comparison of the sensitivity and specificity of the averaging, median, and hyperbolic FDR methods, we annotated gene pairs as previously known synthetic genetic interactors or as not known, based on BioGRID release 2.0.20 (Stark et al. 2006). We then generated receiver operating characteristic (ROC) curves for each method considering all unknown interactions false (Figure 8). In the stringent regime, the hyperbolic FDR method outperforms the average or median, with improved sensitivity (true positive fraction) at high specificity (low false positive fraction). The ROC treats all unknown pairs as false; this underestimates sensitivity since some unknown pairs will actually be true. However, if we assume the impact of false negatives is similar across the methods, our conclusion about the relative ability of each method should hold.

Since the database is not complete, we tested unknown interactions within the 100 top predictions from each method. Some predicted interactions were not tested due to technical problems, such as failure of the PCR for production of the targeting construct. We performed random spore analysis using procedures as described in Pan et al. (2007). As with the ROC based on BioGRID interactions, the hyperbolic FDR provides the best performance, with 72 of 88 tested predictions confirmed or previously known (82%), versus 44 of 82 (54%) for the average and 18 of 86 (21%) for the median.

Using the estimated FDR, we predict 1690 synthetic lethal interactions with FDR cutoff of 0.5, 280 of which are listed in BioGRID (we also predict 22 synthetic rescue

58

interactions). Predicted synthetic lethal and rescue interactions are listed in Tables 2 and 3, respectively. If the FDR is an accurate estimation of false positives, we expect 845 of these predicted synthetic lethal interactions to be true.

Discussion

With 707 out of ~5000 queries yielding at least 845 estimated true interactions from bidirectional data, the number of true synthetic lethal interactions expected per gene is approximately 8. This number is significantly lower than the 34 interactions per query found by Tong et al. (2004). However, the Tong et al. data set contained queries chosen for biological interest, and there were ~30 queries that were abandoned when few potential interactions were detected in the first screen. For both of these reasons the reported data are likely biased toward more interactions per query. Regardless, the true number of synthetic interactions per gene in *S. cerevisiae* is probably more than 8.

When researchers have a particular interest in one of the genes used as a query, additional interactions may be discovered by examining the entire data set, rather than the subset for which bidirectional data are available. The candidates from such a method will contain a higher fraction of false positives, but with verification by random spore or tetrad analysis (Tong et al. 2001; Pan et al. 2007) many additional true positives could be revealed. We present a conservative method for generating synthetic interaction predictions without subsequent verification. Without applying this conservative bidirectional approach, we would predict 11 801 candidate interactions at combined Z-score ≥ 5, or 24 996 at combined Z-score ≥ 4. While candidate interactions may be interesting for selected genes, we focus on expanding the number of dSLAM screens

available for bidirectional analysis, rather than manual verification of these candidates.

Acknowledgments

References

Bolstad BM, Irizarry RA, Astrand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19(2):185–93.

Dobzhansky T. 1946. Genetics of natural populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura*. *Genetics* 31(3):269–90.

Durbin BP, Hardin JS, Hawkins DM, Rocke DM. 2002. A variance-stabilizing transformation for gene-expression microarray data. *Bioinformatics* 18 Suppl 1:S105–10.

Hartman JL 4th, Garvik B, Hartwell L. 2001. Principles for the buffering of genetic variation. *Science* 291(5506):1001–4.

Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential

expression. *Bioinformatics* 18 Suppl 1:S96–104.

Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. *J Comput Graph Stat* 5(3):299–314.

Kelley R, Ideker T. 2005. Systematic interpretation of genetic interactions using protein networks. *Nat Biotechnol* 23(5):561–6.

Ooi SL, Shoemaker DD, Boeke JD 2003. DNA helicase gene interaction network defined using synthetic lethality analyzed by microarray. *Nat Genet* 35(3):277–286

Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD 2004. A robust toolkit for functional profiling of the yeast genome. *Mol Cell* 16(3):487–496

Pan X, Yuan DS, Ooi SL, Wang X, Sookhai-Mahadeo S, Meluh P, Boeke JD 2007. dSLAM analysis of genome-wide genetic interactions in *Saccharomyces cerevisiae*. *Methods* 41(2):206–21.

Peyser BD, Irizarry RA, Tiffany CW, Chen O, Yuan DS, Boeke JD, Spencer FA. Improved statistical analysis of budding yeast TAG microarrays revealed by defined spike-in pools. *Nucleic Acids Res* 33(16):e140.

Schuldiner M, Collins SR, Thompson NJ, Denic V, Bhamidipati A, Punna T, Ihmels J, Andrews B, Boone C, Greenblatt JF, et al. 2005. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* 123(3):507–19.

Shoemaker DD, Lashkari DA, Morris D, Mittmann M, Davis RW. 1996.

Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* 14(4):450–6.

Smyth GK. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3.

Smyth GK. 2005. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, editors. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer. p 397–420.

Smyth GK, Speed TP. 2003. Normalization of cDNA microarray data. *Methods* 31(4): 265–73.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. 2006. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* 34(Database issue):D535–9.

Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, et al. 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294(5550):2364–8.

Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, et al. 2004. Global mapping of the yeast genetic interaction network. *Science* 303(5659):808–13.

Tucker CL, Fields S. 2003. Lethal combinations. *Nat Genet* 35(3):204–5.

Wagner A. 2005. Distributed robustness versus redundancy as causes of mutational robustness. *Bioessays* 27(2):176–88.

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, et al. 1999. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901–6.

Yuan DS, Pan X, Ooi SL, Peyser BD, Spencer FA, Irizarry RA, Boeke JD. 2005. Improved microarray methods for profiling the Yeast Knockout strain collection. *Nucleic Acids Res* 33(12):e103.
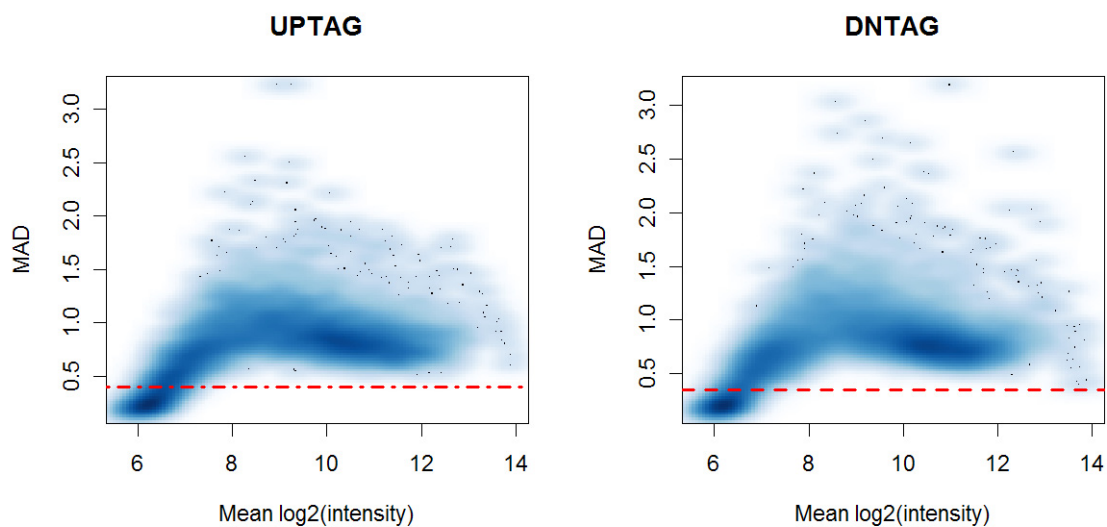
Figure 1. Density scatter plots for UPTAG and DNTAG signal deviation versus intensity. For each UPTAG and DNTAG, the SD of $\log_2$ intensity was estimated using the MAD, and plotted versus the average $\log_2$ intensity for that TAG. Dark blue indicates high density of points. Red lines at 0.40 for UPTAG and 0.35 for DNTAG indicate MAD cutoffs for annotating failed TAGs.

TAG Behavior
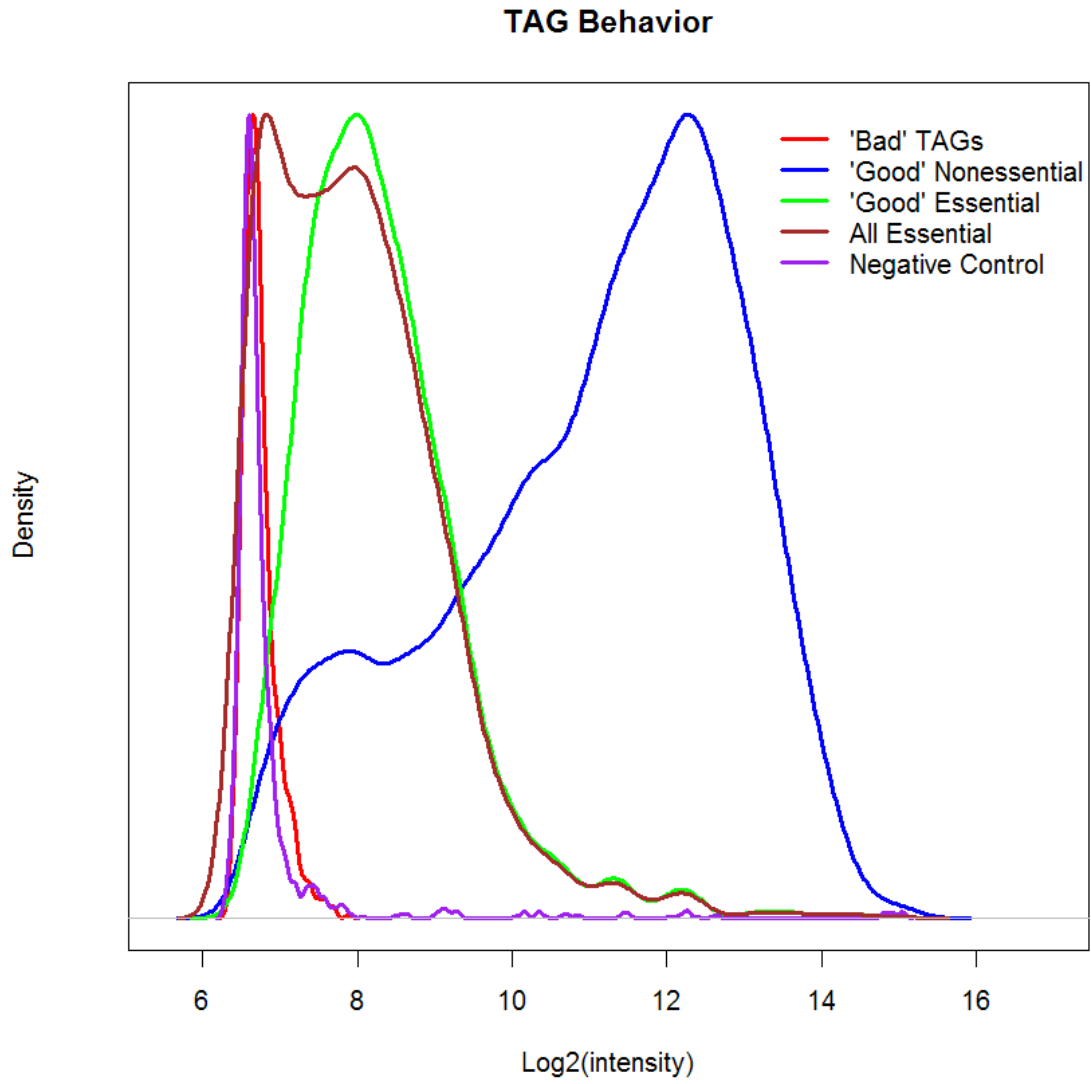
Figure 2. TAG intensity by type. Density histograms showing probability densities (*y*-axis) for average $\log_2$ intensities (*x*-axis) of various categories of TAGs across 309 microarrays. Both UPTAG and DNTAG values are included in each category.

Figure 3. Behavior across 309 arrays for 2 UPTAGs. (A) Density histograms showing probability densities (y-axis) for $\log_2$ ratios (x-axis) of 2 UPTAGs across 309 normalized microarrays. All arrays were prepared from a single batch of labeled primer. Blue line is the distribution of values for a typical TAG, and red line is from a TAG that displays primer-batch–specific artifacts. Ticks along the x-axis show location for each value. Rectangles span the mean ±1 SD. (B) Density histograms showing probability densities for the $\log_2$ ratio Z-scores of the 2 UPTAGs shown in (A). Rectangles span the mean (0) ±1 SD. Arrow indicates elevated Z-score in one microarray for the TAG shown in red.

Figure 3.

Behavior of Two TAGs Across 309 Arrays

A.

Behavior of Two TAGs Across 309 Arrays

B.
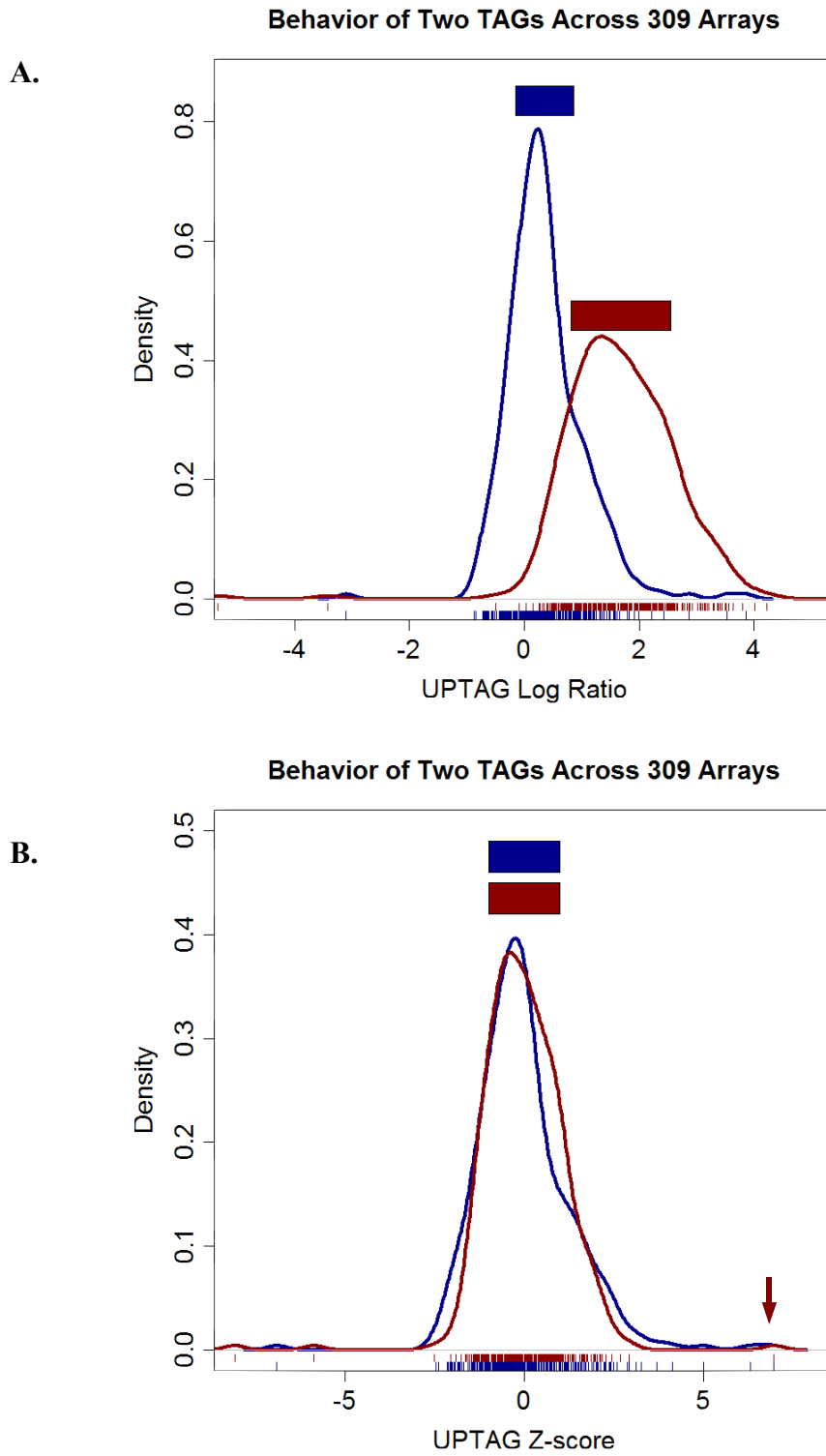
UP− and DNTAGs from one array
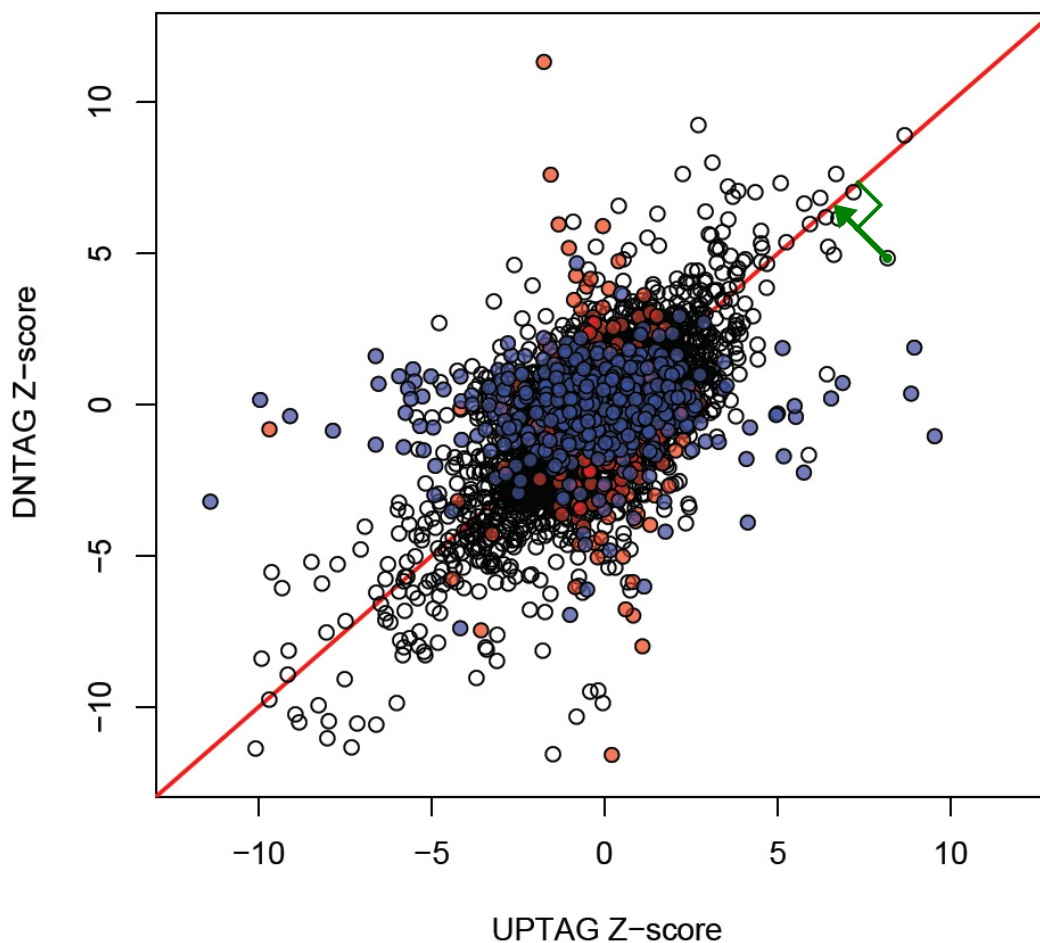
Figure 4. UP- and DNTAGs from a single hybridization. The DNTAG *Z*-score is plotted versus the UPTAG *Z*-score for all YKO strains in a single experiment. Red points represent strains with the UPTAG annotated bad and a working DNTAG, blue points are bad DNTAG and good UPTAG, and open circles are YKOs with two good TAGs. Green arrow represents the weighted average transformation for a point with two good TAGs.

Biological replicate experiments
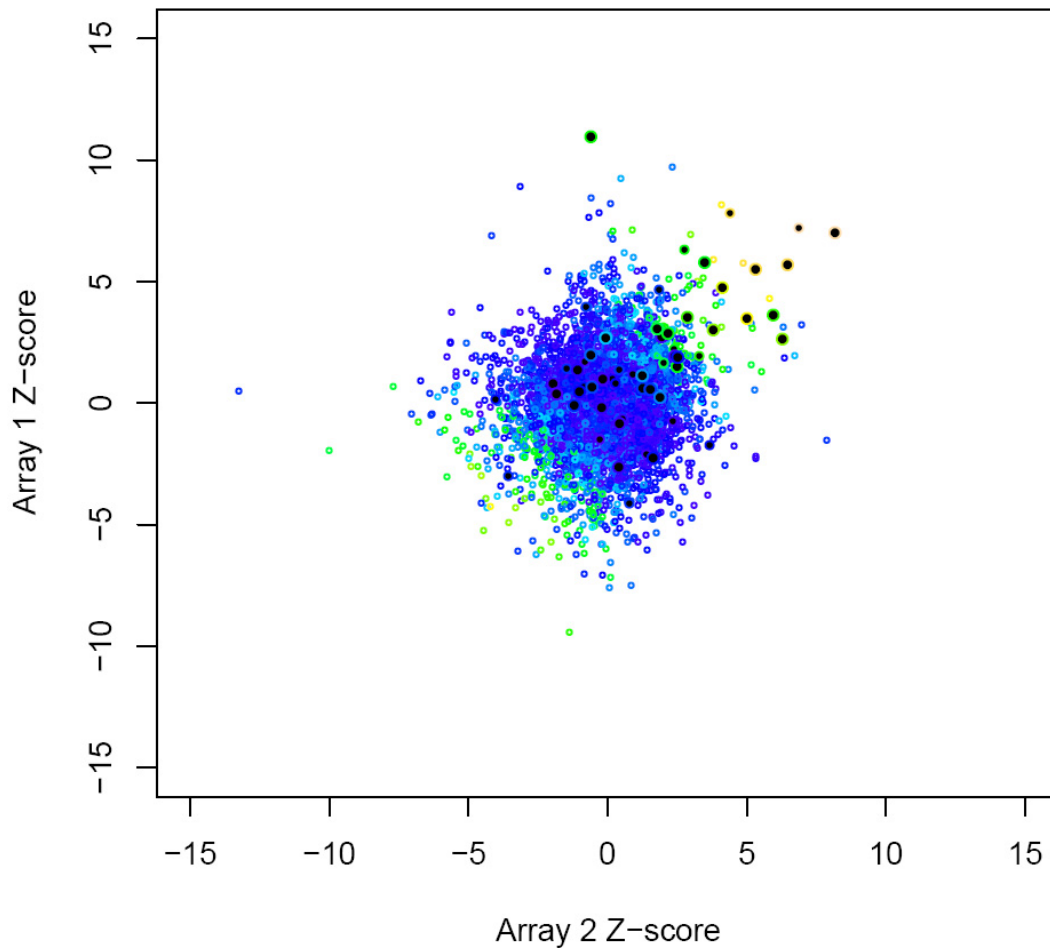
Figure 5. Replicate experiments reveal variability. Three biological replicate experiments were performed with *YCL016C* (*DCC1*), and *p*-values were generated using an empirical Bayes procedure. Combined *Z*-scores from two of the arrays for each YKO are shown, with circle color representing *p*-value on a logarithmic curve from $p = 0$ (warm colors: orange) to $p = 1$ (cool colors: blue). Filled black circles are known synthetic lethal (large circles) or fitness defect (small circles) from Tong et al. (2004).
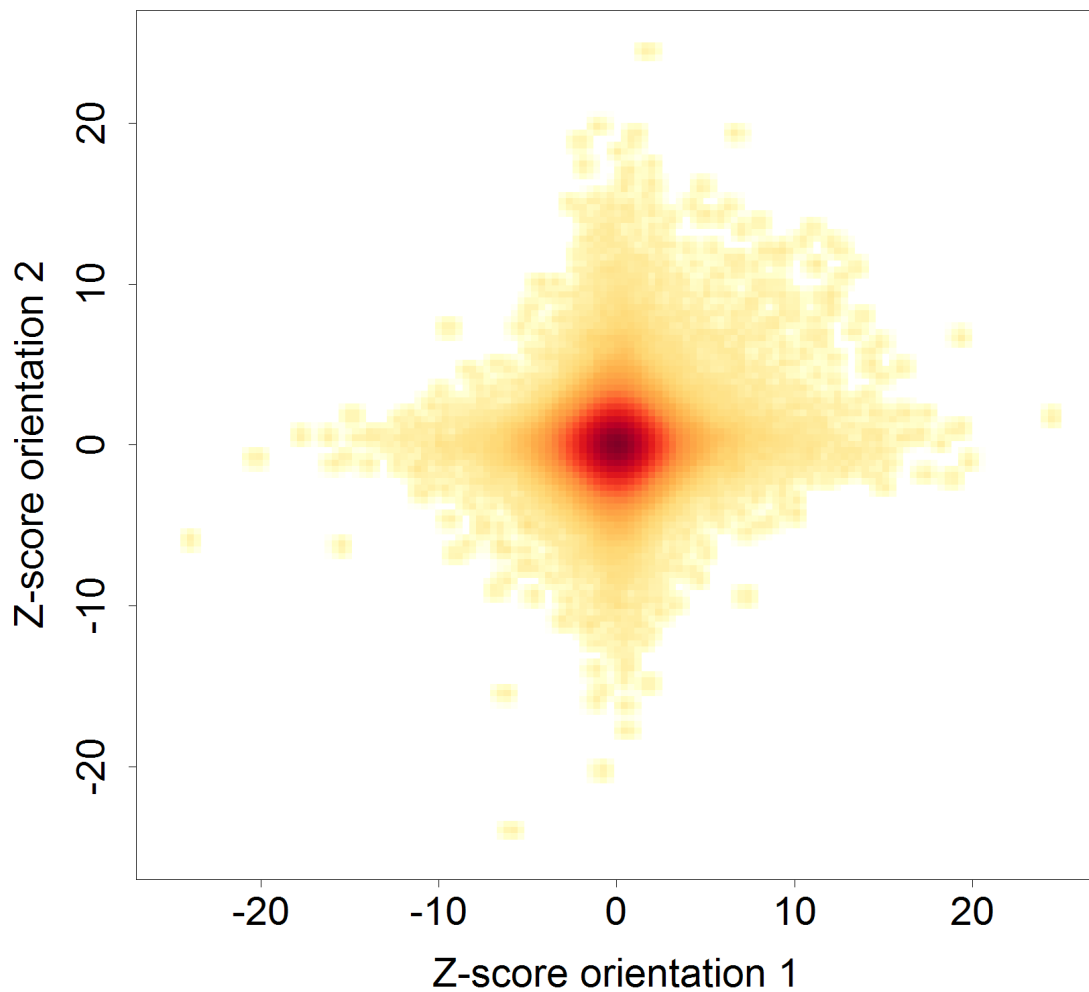
Figure 6. *Z*-score agreement suggests true interactions. Density scatter plot of combined *Z*-scores of two marker orientations for each gene pair. Darker orange regions represent higher density of points. More points are found in the upper right quadrant and lower left quadrant, where *Z*-scores agree between two marker orientations, than in the lower right and upper left quadrants.
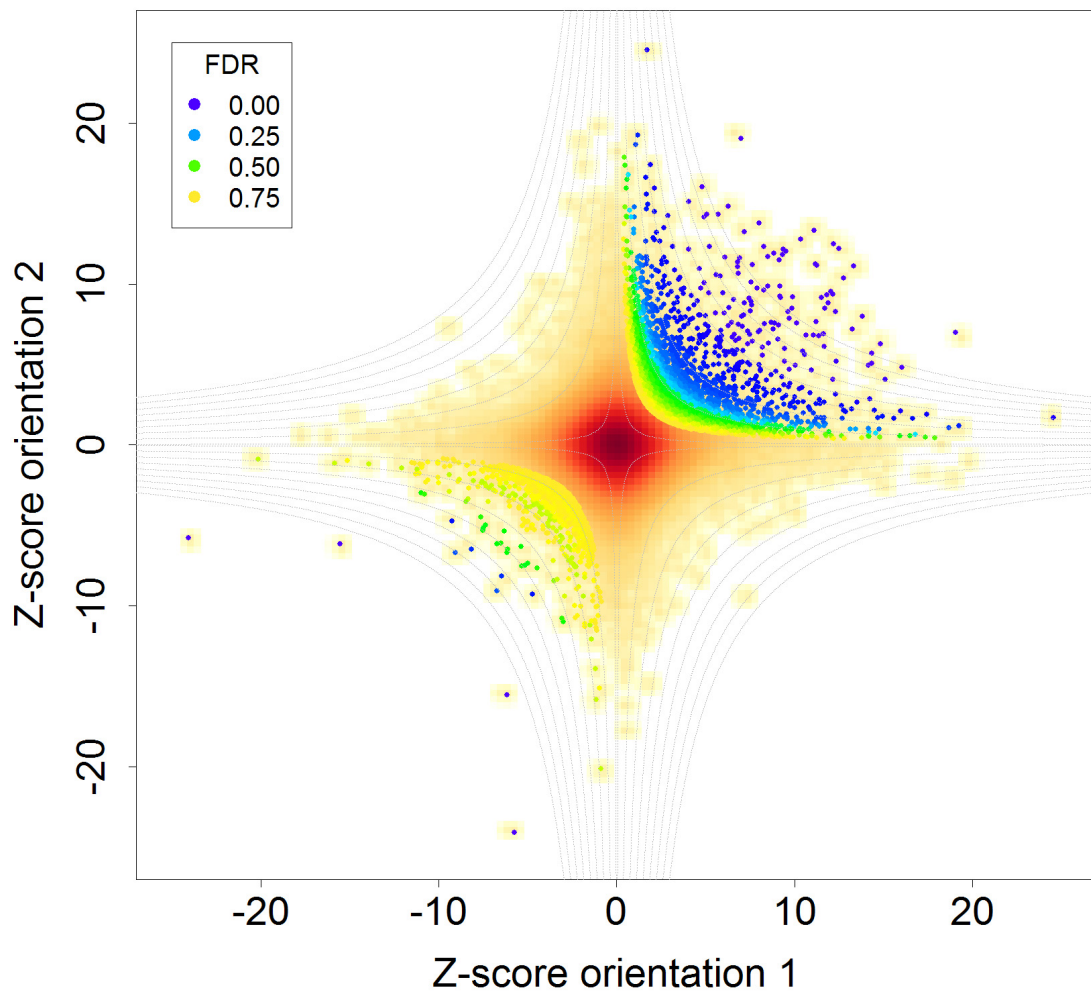
Figure 7. Estimation of noise using distribution of *Z*-scores. Density scatter plot of combined *Z*-scores for two marker orientations from Figure 6 overlain with estimated FDR. Selected hyperbolic cutoffs are shown in gray. Points are shown and colored by estimated FDR for all gene pairs up to FDR = 0.75. Both synthetic lethal (upper right) and synthetic rescue (lower left) are shown.

Figure 8. ROC curves. Three methods for combining information from two marker orientations were compared for specificity and sensitivity using synthetic genetic interactions found in BioGRID release 2.0.20 as true. All other interactions were considered false. The true positive fraction was plotted versus the false positive fraction for decreasing average or median Z-score values, and increasing FDR values. (A) The segment of the ROC curve showing up to ~1000 false positive interactions and ~200 true positive interactions. (B) and (C) Expanded segments of the ROC curves.

Figure 8.

A.

B.   **C.**

73
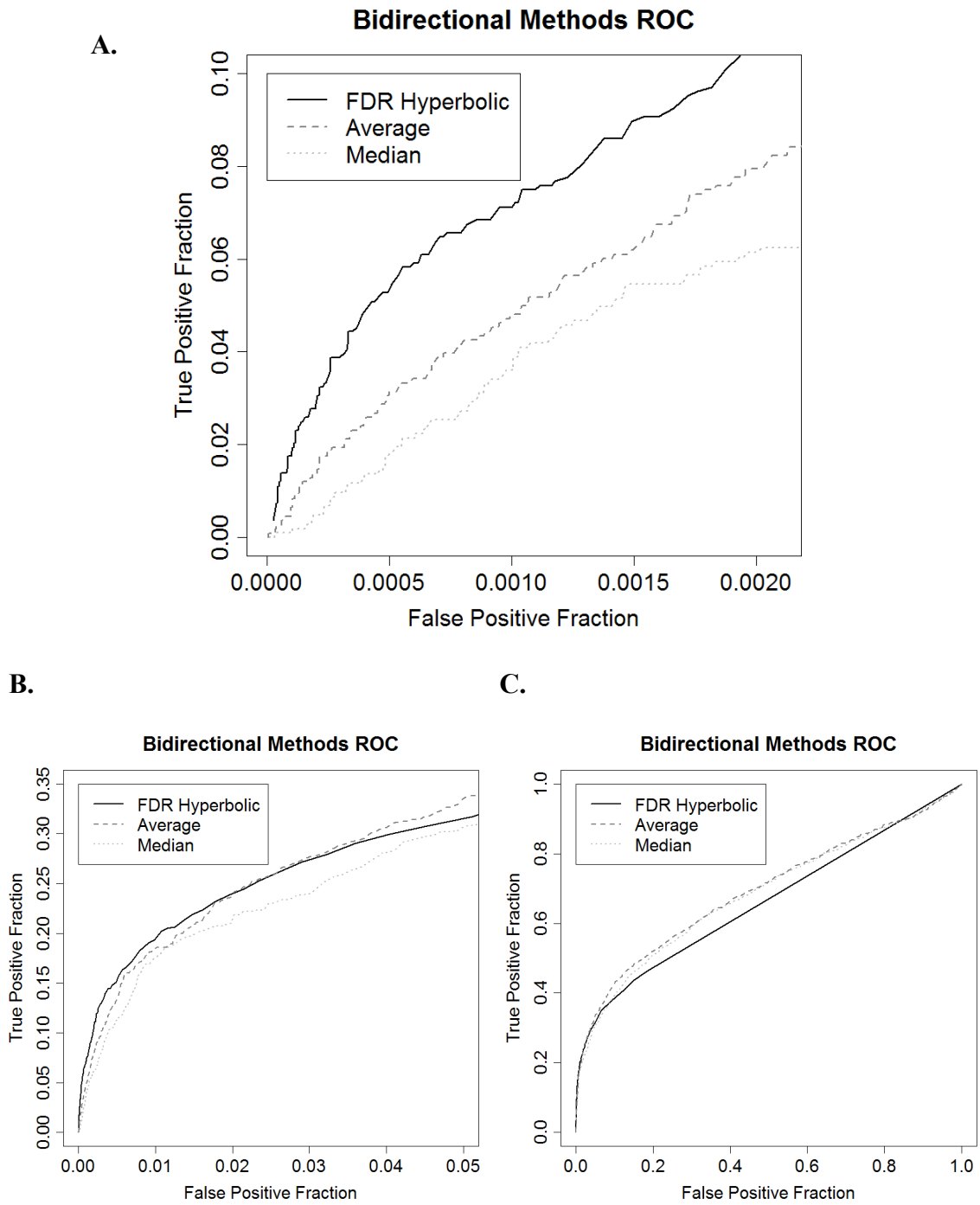
Table 1. Strains responsive to presence of uracil.

Strain Location	Strain ID	ORF	Gene Symbol	Function
Poor growth on -uracil media	202E9	21569	YLR014C	PPR1	positively regulates transcription of genes involved in uracil biosynthesis
	202E10	21570	YLR015W	BRE2	~250 bp upstream of PPR1
	202F10	21582	YLR027C	AAT2	aspartate biosynthesis: aspartate aminotransferase
	219E1	24081	YLR420W	URA4	de novo biosynthesis of pyrimidines: dihydroorotase
	225F8	21295	YJL130C	URA2	de novo biosynthesis of pyrimidines: aspartate transcarbamylase
	235D11	25066	YKL216W	URA1	de novo biosynthesis of pyrimidines: dihydroorotate dehydrogenase
	241C4	26506	YML106W	URA5	de novo biosynthesis of pyrimidines: orotate phosphoribosyltransferase 1
	241C5	26507	YML107C	YML107C	~500 bp upstream of URA5
Poor growth on +uracil media	208C1	21598	YOR302W	YOR302W	regulates translation of the CPA1 mRNA
	208C2	21599	YOR303W	CPA1	carbamoyl phosphate synthetase
	245G10	26916	YJR109C	CPA2	carbamyl phosphate synthetase
	253B4	23158	YBR021W	FUR4	uracil permease
	256G2	27336	YOR302W	YOR302W	regulates translation of the CPA1 mRNA
	256G3	27337	YOR303W	CPA1	carbamoyl phosphate synthetase

Table 2. Predicted synthetic lethal pairs by estimated FDR.

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YBR223C	YHR134W	0	YMR263W	YLR418C	0.040323
YBR224W	YHR134W	0	YBR219C	YLR023C	0.041667
YBR228W	YMR190C	0	YLR023C	YBR219C	0.041667
YBR229C	YNL322C	0	YAL029C	YOR054C	0.052326
YDR072C	YDR372C	0	YOR054C	YAL029C	0.052326
YDR372C	YDR072C	0	YBR215W	YLR418C	0.052632
YDR388W	YLR111W	0	YLR418C	YBR215W	0.052632
YDR399W	YGR061C	0	YMR078C	YOR026W	0.052632
YER056C	YGR061C	0	YOR026W	YMR078C	0.052632
YGR061C	YDR399W	0	YLR125W	YPL072W	0.053571
YGR061C	YER056C	0	YPL072W	YLR125W	0.053571
YHR134W	YBR223C	0	YDR117C	YNL298W	0.054054
YHR134W	YBR224W	0	YNL298W	YDR117C	0.054054
YLR102C	YNL298W	0	YGR188C	YOR066W	0.054348
YLR111W	YDR388W	0	YOR066W	YGR188C	0.054348
YMR190C	YBR228W	0	YDR389W	YNR051C	0.054688
YMR238W	YNL322C	0	YNR051C	YDR389W	0.054688
YNL298W	YLR102C	0	YER116C	YNL077W	0.054945
YNL322C	YBR229C	0	YNL077W	YER116C	0.054945
YNL322C	YMR238W	0	YDL101C	YNL281W	0.055556
YOR035C	YOR039W	0	YLR418C	YOL004W	0.055556
YOR039W	YOR035C	0	YNL281W	YDL101C	0.055556
YLR418C	YOR038C	0.008772	YOL004W	YLR418C	0.055556
YOR038C	YLR418C	0.008772	YMR078C	YNL273W	0.05618
YOR026W	YPR120C	0.009091	YNL273W	YMR078C	0.05618
YPR120C	YOR026W	0.009091	YBR009C	YBR215W	0.05625
YDR414C	YNL322C	0.009259	YBR215W	YBR009C	0.05625
YNL322C	YDR414C	0.009259	YCL016C	YNL273W	0.05625
YNL298W	YPL161C	0.009615	YNL273W	YCL016C	0.05625
YPL161C	YNL298W	0.009615	YNL273W	YPR135W	0.056338
YOL044W	YOR322C	0.01	YPR135W	YNL273W	0.056338
YOR028C	YOR033C	0.01	YDR414C	YNL297C	0.057471
YOR033C	YOR028C	0.01	YNL297C	YDR414C	0.057471
YOR322C	YOL044W	0.01	YBR200W	YLL021W	0.057971
YLR089C	YNL277W	0.01087	YEL037C	YML013C-A	0.057971
YNL277W	YLR089C	0.01087	YLL021W	YBR200W	0.057971
YNL298W	YOL004W	0.011364	YML013C-A	YEL037C	0.057971
YOL004W	YNL298W	0.011364	YDR399W	YDR408C	0.059701
YER116C	YMR190C	0.011628	YDR408C	YDR399W	0.059701
YMR190C	YER116C	0.011628	YDL013W	YMR190C	0.0625
YHR191C	YPR135W	0.011905	YMR190C	YDL013W	0.0625
YPR135W	YHR191C	0.011905	YMR186W	YPL161C	0.063158
YLL021W	YNL298W	0.013514	YPL161C	YMR186W	0.063158
YNL298W	YLL021W	0.013514	YER016W	YHR191C	0.06383
YDR388W	YOL004W	0.013889	YHR191C	YER016W	0.06383
YOL004W	YDR388W	0.013889	YLR023C	YOR067C	0.064516
YBR009C	YOR038C	0.014706	YOR067C	YLR023C	0.064516
YOR038C	YBR009C	0.014706	YKL113C	YMR224C	0.070093
YLR418C	YMR263W	0.040323	YMR224C	YKL113C	0.070093

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YHR015W	YPL248C	0.071429	YOR033C	YOR054C	0.085821
YMR246W	YPL100W	0.071429	YOR054C	YOR033C	0.085821
YPL100W	YMR246W	0.071429	YDR114C	YPL139C	0.086466
YPL248C	YHR015W	0.071429	YGR188C	YPL253C	0.086466
YMR263W	YNL298W	0.072072	YOR039W	YOR071C	0.086466
YNL298W	YMR263W	0.072072	YOR071C	YOR039W	0.086466
YBR231C	YLR418C	0.072816	YPL139C	YDR114C	0.086466
YDR388W	YML094W	0.072816	YPL253C	YGR188C	0.086466
YLR418C	YBR231C	0.072816	YLR079W	YOR346W	0.087413
YML094W	YDR388W	0.072816	YOR346W	YLR079W	0.087413
YFR036W	YNL298W	0.074257	YEL043W	YMR246W	0.088028
YNL298W	YFR036W	0.074257	YMR246W	YEL043W	0.088028
YFR036W	YNR051C	0.075	YDR372C	YNR051C	0.088235
YNR051C	YFR036W	0.075	YER016W	YPL155C	0.088235
YGR188C	YLR079W	0.075221	YNR051C	YDR372C	0.088235
YLR079W	YGR188C	0.075221	YPL155C	YER016W	0.088235
YFR036W	YPL106C	0.075893	YML094W	YOR026W	0.088652
YPL106C	YFR036W	0.075893	YOR026W	YML094W	0.088652
YDR109C	YHR015W	0.076531	YMR228W	YPL133C	0.089286
YHR015W	YDR109C	0.076531	YPL133C	YMR228W	0.089286
YLR111W	YNR051C	0.07732	YBR217W	YMR246W	0.09058
YNR051C	YLR111W	0.07732	YMR246W	YBR217W	0.09058
YDL101C	YLR418C	0.077586	YCL016C	YPR141C	0.090604
YLR418C	YDL101C	0.077586	YPR141C	YCL016C	0.090604
YFR036W	YLR079W	0.078261	YDR076W	YKL113C	0.091241
YLR079W	YFR036W	0.078261	YKL113C	YDR076W	0.091241
YNL281W	YNL291C	0.078261	YLR023C	YOR313C	0.106667
YNL291C	YNL281W	0.078261	YOR313C	YLR023C	0.106667
YAL039C	YPL270W	0.079365	YDR093W	YMR253C	0.108553
YPL270W	YAL039C	0.079365	YMR253C	YDR093W	0.108553
YLR102C	YMR198W	0.08	YMR198W	YPR120C	0.109272
YMR048W	YMR078C	0.08	YPR120C	YMR198W	0.109272
YMR078C	YMR048W	0.08	YGR188C	YHR191C	0.113924
YMR198W	YLR102C	0.08	YHR191C	YGR188C	0.113924
YMR282C	YOR065W	0.082031	YML010W-A	YMR263W	0.113924
YOR065W	YMR282C	0.082031	YMR263W	YML010W-A	0.113924
YOR011W	YOR026W	0.082645	YNL021W	YNL294C	0.113924
YOR026W	YOR011W	0.082645	YNL294C	YNL021W	0.113924
YML010W-A	YOR038C	0.083333	YDR359C	YMR198W	0.116129
YOR038C	YML010W-A	0.083333	YMR198W	YDR359C	0.116129
YBL052C	YMR246W	0.084034	YDR096W	YMR194W	0.11747
YBR009C	YOR026W	0.084034	YML079W	YNL288W	0.11747
YEL017W	YEL023C	0.084034	YMR194W	YDR096W	0.11747
YEL023C	YEL017W	0.084034	YNL288W	YML079W	0.11747
YMR246W	YBL052C	0.084034	YCL016C	YGR188C	0.117647
YOR026W	YBR009C	0.084034	YGR188C	YCL016C	0.117647
YGL163C	YKL113C	0.084615	YHR191C	YOR026W	0.11875
YKL113C	YGL163C	0.084615	YOR026W	YHR191C	0.11875
YLR079W	YOR026W	0.084615	YBR231C	YDL074C	0.118902
YOR026W	YLR079W	0.084615	YDL074C	YBR231C	0.118902
YLR125W	YML089C	0.085616	YLR072W	YLR082C	0.118902
YML089C	YLR125W	0.085616	YLR082C	YLR072W	0.118902

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YMR183C	YPL232W	0.118902	YDL074C	YMR198W	0.1425
YOR011W	YOR035C	0.118902	YEL024W	YMR282C	0.1425
YOR035C	YOR011W	0.118902	YMR198W	YDL074C	0.1425
YPL232W	YMR183C	0.118902	YMR198W	YNL273W	0.1425
YCL016C	YMR048W	0.119186	YMR282C	YEL024W	0.1425
YCL061C	YMR198W	0.119186	YNL273W	YMR198W	0.1425
YGR188C	YNL273W	0.119186	YBR009C	YMR179W	0.144841
YMR048W	YCL016C	0.119186	YDL074C	YPL254W	0.144841
YMR198W	YCL061C	0.119186	YLL015W	YML058C-A	0.144841
YNL273W	YGR188C	0.119186	YML058C-A	YLL015W	0.144841
YFR036W	YMR198W	0.121302	YMR179W	YBR009C	0.144841
YKR029C	YMR194W	0.121302	YMR179W	YPL102C	0.144841
YMR194W	YKR029C	0.121302	YPL102C	YMR179W	0.144841
YMR198W	YFR036W	0.121302	YPL254W	YDL074C	0.144841
YDR414C	YMR214W	0.124277	YAL039C	YDR350C	0.145228
YMR214W	YDR414C	0.124277	YBR009C	YFR036W	0.145228
YLR065C	YPL106C	0.132184	YBR107C	YER016W	0.145228
YPL106C	YLR065C	0.132184	YDR014W	YHR015W	0.145228
YMR233W	YOR054C	0.132768	YDR350C	YAL039C	0.145228
YOR054C	YMR233W	0.132768	YER016W	YBR107C	0.145228
YER116C	YOR156C	0.134286	YFR036W	YBR009C	0.145228
YOR156C	YER116C	0.134286	YHR015W	YDR014W	0.145228
YLR125W	YPL136W	0.138743	YLR125W	YOR084W	0.145228
YMR273C	YOR011W	0.138743	YOR054C	YPL107W	0.145228
YNL303W	YOR054C	0.138743	YOR084W	YLR125W	0.145228
YOR011W	YMR273C	0.138743	YPL107W	YOR054C	0.145228
YOR054C	YNL303W	0.138743	YCL061C	YHR031C	0.146341
YPL136W	YLR125W	0.138743	YHR031C	YCL061C	0.146341
YMR283C	YNL269W	0.139037	YJR043C	YNR051C	0.146341
YNL269W	YMR283C	0.139037	YNL294C	YNL298W	0.146341
YOL020W	YOR002W	0.139037	YNL298W	YNL294C	0.146341
YOR002W	YOL020W	0.139037	YNR051C	YJR043C	0.146341
YBR231C	YMR263W	0.139594	YOR026W	YPL253C	0.146341
YEL001C	YPL115C	0.139594	YPL253C	YOR026W	0.146341
YMR263W	YBR231C	0.139594	YDR134C	YLR023C	0.146809
YNL297C	YPL120W	0.139594	YDR359C	YOR024W	0.146809
YNL335W	YPL125W	0.139594	YLR023C	YDR134C	0.146809
YPL115C	YEL001C	0.139594	YNL269W	YOL011W	0.146809
YPL120W	YNL297C	0.139594	YOL011W	YNL269W	0.146809
YPL125W	YNL335W	0.139594	YOR024W	YDR359C	0.146809
YLL016W	YLR093C	0.14011	YPL118W	YPL133C	0.146809
YLR093C	YLL016W	0.14011	YPL133C	YPL118W	0.146809
YOR026W	YOR080W	0.14011	YMR012W	YPL103C	0.147196
YOR054C	YOR071C	0.14011	YNL294C	YNR051C	0.147196
YOR071C	YOR054C	0.14011	YNR051C	YNL294C	0.147196
YOR080W	YOR026W	0.14011	YOL076W	YPR135W	0.147196
YBR215W	YML010W-A	0.140625	YPL103C	YMR012W	0.147196
YML010W-A	YBR215W	0.140625	YPR135W	YOL076W	0.147196
YER016W	YOL004W	0.141304	YDR393W	YOR049C	0.147343
YOL004W	YER016W	0.141304	YEL031W	YHR012W	0.147343
YHR031C	YJL047C	0.142077	YFR036W	YNL307C	0.147343
YJL047C	YHR031C	0.142077	YGL163C	YOR144C	0.147343

77

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YHR012W	YEL031W	0.147343	YDL101C	YOR039W	0.163462
YNL297C	YOR068C	0.147343	YDR093W	YOR005C	0.163462
YNL307C	YFR036W	0.147343	YDR191W	YOR024W	0.163462
YOR049C	YDR393W	0.147343	YEL001C	YEL008W	0.163462
YOR068C	YNL297C	0.147343	YEL008W	YEL001C	0.163462
YOR144C	YGL163C	0.147343	YEL053C	YNL297C	0.163462
YHR191C	YMR198W	0.147465	YMR078C	YPR135W	0.163462
YMR198W	YHR191C	0.147465	YNL297C	YEL053C	0.163462
YOR037W	YPL260W	0.147465	YOR005C	YDR093W	0.163462
YPL260W	YOR037W	0.147465	YOR024W	YDR191W	0.163462
YGR188C	YPR120C	0.147727	YOR039W	YDL101C	0.163462
YHR015W	YPL105C	0.147727	YPR135W	YMR078C	0.163462
YHR110W	YHR134W	0.147727	YBR213W	YOR054C	0.168519
YHR134W	YHR110W	0.147727	YDR101C	YLR018C	0.168519
YPL105C	YHR015W	0.147727	YDR389W	YML013C-A	0.168519
YPR120C	YGR188C	0.147727	YEL053C	YLR023C	0.168519
YBR201W	YOR040W	0.148515	YGR188C	YMR078C	0.168519
YDR388W	YOR035C	0.148515	YLR018C	YDR101C	0.168519
YOR035C	YDR388W	0.148515	YLR023C	YEL053C	0.168519
YOR040W	YBR201W	0.148515	YLR111W	YNL297C	0.168519
YKR029C	YOL004W	0.149289	YLR418C	YOR308C	0.168519
YLL015W	YMR282C	0.149289	YML013C-A	YDR389W	0.168519
YLL016W	YOR054C	0.149289	YMR078C	YGR188C	0.168519
YLR023C	YMR163C	0.149289	YNL277W	YNL303W	0.168519
YMR163C	YLR023C	0.149289	YNL297C	YLR111W	0.168519
YMR282C	YLL015W	0.149289	YNL303W	YNL277W	0.168519
YOL004W	YKR029C	0.149289	YOR054C	YBR213W	0.168519
YOR054C	YLL016W	0.149289	YOR308C	YLR418C	0.168519
YBL052C	YPL125W	0.149351	YBR173C	YOL004W	0.170139
YDR334W	YGR188C	0.149351	YDL101C	YDR435C	0.170139
YEL023C	YOR005C	0.149351	YDL101C	YPL152W	0.170139
YGL194C	YOL004W	0.149351	YDR049W	YML013C-A	0.170139
YGR188C	YDR334W	0.149351	YDR435C	YDL101C	0.170139
YJR043C	YNL273W	0.149351	YKR029C	YOR037W	0.170139
YNL273W	YJR043C	0.149351	YML013C-A	YDR049W	0.170139
YOL004W	YGL194C	0.149351	YNL277W	YOL001W	0.170139
YOR005C	YEL023C	0.149351	YOL001W	YNL277W	0.170139
YOR034C	YOR037W	0.149351	YOL004W	YBR173C	0.170139
YOR037W	YOR034C	0.149351	YOR037W	YKR029C	0.170139
YPL116W	YPL119C	0.149351	YPL152W	YDL101C	0.170139
YPL119C	YPL116W	0.149351	YLR065C	YLR125W	0.170956
YPL125W	YBL052C	0.149351	YLR125W	YLR065C	0.170956
YDR408C	YER056C	0.151786	YBR215W	YOR080W	0.172598
YEL040W	YEL049W	0.151786	YBR266C	YDR083W	0.172598
YEL049W	YEL040W	0.151786	YDR083W	YBR266C	0.172598
YER056C	YDR408C	0.151786	YHR109W	YOR315W	0.172598
YLR021W	YML013C-A	0.151786	YML010W-A	YNR051C	0.172598
YML013C-A	YLR021W	0.151786	YNR051C	YML010W-A	0.172598
YMR198W	YNL307C	0.151786	YOR080W	YBR215W	0.172598
YNL307C	YMR198W	0.151786	YOR315W	YHR109W	0.172598
YMR230W	YNL297C	0.162835	YBR194W	YKR029C	0.174545
YNL297C	YMR230W	0.162835	YDR108W	YJR060W	0.174545

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YJR060W	YDR108W	0.174545	YDR116C	YOR009W	0.197947
YKR029C	YBR194W	0.174545	YDR353W	YDR096W	0.197947
YOR054C	YOR383C	0.174545	YEL010W	YEL043W	0.197947
YOR383C	YOR054C	0.174545	YEL043W	YEL010W	0.197947
YDR097C	YLR023C	0.178571	YGL058W	YBR200W	0.197947
YDR388W	YHR111W	0.178571	YLR097C	YLR111W	0.197947
YHR012W	YOR322C	0.178571	YLR111W	YLR097C	0.197947
YHR111W	YDR388W	0.178571	YML010W-A	YMR179W	0.197947
YLR023C	YDR097C	0.178571	YML094W	YBR231C	0.197947
YLR079W	YOR025W	0.178571	YML094W	YCL016C	0.197947
YLR125W	YOL054W	0.178571	YMR179W	YML010W-A	0.197947
YOL054W	YLR125W	0.178571	YNL307C	YPR120C	0.197947
YOR025W	YLR079W	0.178571	YOL001W	YOR002W	0.197947
YOR322C	YHR012W	0.178571	YOL050C	YOL054W	0.197947
YDR097C	YPL118W	0.182274	YOL054W	YOL050C	0.197947
YDR388W	YNL293W	0.182274	YOR002W	YOL001W	0.197947
YHR134W	YOR054C	0.182274	YOR009W	YDR116C	0.197947
YMR257C	YMR282C	0.182274	YPR120C	YNL307C	0.197947
YMR282C	YMR257C	0.182274	YBL058W	YER060W	0.198171
YNL293W	YDR388W	0.182274	YCL061C	YJR043C	0.198171
YNL335W	YOR038C	0.182274	YDR114C	YKR029C	0.198171
YOR038C	YNL335W	0.182274	YER060W	YBL058W	0.198171
YOR054C	YHR134W	0.182274	YJR043C	YCL061C	0.198171
YPL118W	YDR097C	0.182274	YKR029C	YDR114C	0.198171
YBR195C	YOR038C	0.195724	YLL016W	YMR273C	0.198171
YEL031W	YLR418C	0.195724	YMR273C	YLL016W	0.198171
YEL040W	YMR273C	0.195724	YDR399W	YMR120C	0.198697
YER095W	YOR011W	0.195724	YMR120C	YDR399W	0.198697
YLR015W	YLR065C	0.195724	YDR372C	YOL001W	0.199128
YLR065C	YLR015W	0.195724	YDR392W	YPL129W	0.199128
YLR418C	YEL031W	0.195724	YEL037C	YEL040W	0.199128
YMR273C	YEL040W	0.195724	YEL040W	YEL037C	0.199128
YOR011W	YER095W	0.195724	YOL001W	YDR372C	0.199128
YOR038C	YBR195C	0.195724	YPL129W	YDR392W	0.199128
YBR194W	YGL194C	0.196203	YBR231C	YLR125W	0.200935
YBR200W	YLR023C	0.196203	YBR255W	YOR035C	0.200935
YDR414C	YPL120W	0.196203	YDR063W	YLR023C	0.200935
YGL194C	YBR194W	0.196203	YDR094W	YLR125W	0.200935
YLL063C	YLR023C	0.196203	YLR023C	YDR063W	0.200935
YLR023C	YBR200W	0.196203	YLR125W	YBR231C	0.200935
YLR023C	YLL063C	0.196203	YLR125W	YDR094W	0.200935
YMR198W	YNL330C	0.196203	YMR048W	YMR198W	0.200935
YNL297C	YPL133C	0.196203	YMR198W	YMR048W	0.200935
YNL298W	YPL254W	0.196203	YOR035C	YBR255W	0.200935
YNL330C	YMR198W	0.196203	YBR180W	YEL053C	0.206553
YPL120W	YDR414C	0.196203	YEL053C	YBR180W	0.206553
YPL133C	YNL297C	0.196203	YLR028C	YMR120C	0.206553
YPL254W	YNL298W	0.196203	YML008C	YNL322C	0.206553
YBR200W	YGL058W	0.197947	YMR120C	YLR028C	0.206553
YBR231C	YML094W	0.197947	YNL298W	YPL106C	0.206553
YCL016C	YML094W	0.197947	YNL322C	YML008C	0.206553
YDR096W	YDR353W	0.197947	YNR032C-A	YOR308C	0.206553

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YOR003W	YOR038C	0.206553	YMR252C	YLR125W	0.220317
YOR038C	YOR003W	0.206553	YNL330C	YAL039C	0.220317
YOR308C	YNR032C-A	0.206553	YOR002W	YOR011W	0.220317
YPL106C	YNL298W	0.206553	YOR011W	YOR002W	0.220317
YPL115C	YPL119C	0.206553	YOR025W	YDR191W	0.220317
YPL119C	YPL115C	0.206553	YOR054C	YMR158W-A	0.220317
YJL047C	YNL303W	0.213277	YDR109C	YLR094C	0.226804
YLR023C	YLR053C	0.213277	YER056C	YMR120C	0.226804
YLR053C	YLR023C	0.213277	YGL194C	YML013W	0.226804
YLR125W	YPL099C	0.213277	YKR029C	YMR193C-A	0.226804
YNL303W	YJL047C	0.213277	YLR015W	YLR023C	0.226804
YPL099C	YLR125W	0.213277	YLR023C	YLR015W	0.226804
YDR372C	YNL297C	0.217877	YLR094C	YDR109C	0.226804
YDR435C	YEL031W	0.217877	YML013W	YGL194C	0.226804
YEL031W	YDR435C	0.217877	YML079W	YOR351C	0.226804
YFR036W	YOR026W	0.217877	YMR120C	YER056C	0.226804
YLR079W	YOR014W	0.217877	YMR186W	YMR194W	0.226804
YNL297C	YDR372C	0.217877	YMR193C-A	YKR029C	0.226804
YOR014W	YLR079W	0.217877	YMR194W	YMR186W	0.226804
YOR026W	YFR036W	0.217877	YMR246W	YMR291W	0.226804
YBR009C	YGR188C	0.219346	YMR291W	YMR246W	0.226804
YDR093W	YOR034C	0.219346	YOR065W	YPL270W	0.226804
YDR350C	YMR256C	0.219346	YOR351C	YML079W	0.226804
YDR388W	YEL003W	0.219346	YPL270W	YOR065W	0.226804
YEL003W	YDR388W	0.219346	YDR414C	YOL001W	0.232824
YEL018W	YER016W	0.219346	YMR214W	YOL023W	0.232824
YEL029C	YMR263W	0.219346	YMR238W	YOR054C	0.232824
YER016W	YEL018W	0.219346	YMR273C	YMR291W	0.232824
YGR188C	YBR009C	0.219346	YMR291W	YMR273C	0.232824
YLL024C	YML059C	0.219346	YNL297C	YOL004W	0.232824
YML059C	YLL024C	0.219346	YOL001W	YDR414C	0.232824
YMR256C	YDR350C	0.219346	YOL004W	YNL297C	0.232824
YMR263W	YEL029C	0.219346	YOL023W	YMR214W	0.232824
YNL322C	YNR051C	0.219346	YOR054C	YMR238W	0.232824
YNR051C	YNL322C	0.219346	YBR231C	YOL004W	0.235222
YOR034C	YDR093W	0.219346	YDL074C	YJL168C	0.235222
YAL039C	YNL330C	0.220317	YDR392W	YML013C-A	0.235222
YBR009C	YML010W-A	0.220317	YFR036W	YHR191C	0.235222
YBR107C	YGR188C	0.220317	YHR191C	YFR036W	0.235222
YBR194W	YEL031W	0.220317	YIR004W	YLR023C	0.235222
YDR073W	YLR125W	0.220317	YJL168C	YDL074C	0.235222
YDR191W	YOR025W	0.220317	YLL055W	YOR054C	0.235222
YDR350C	YDR353W	0.220317	YLR023C	YIR004W	0.235222
YDR353W	YDR350C	0.220317	YML013C-A	YDR392W	0.235222
YEL031W	YBR194W	0.220317	YML013C-A	YNL307C	0.235222
YGR188C	YBR107C	0.220317	YNL307C	YML013C-A	0.235222
YLR125W	YDR073W	0.220317	YOL004W	YBR231C	0.235222
YLR125W	YMR252C	0.220317	YOR013W	YOR028C	0.235222
YML010W-A	YBR009C	0.220317	YOR019W	YOR023C	0.235222
YMR158W-A	YOR054C	0.220317	YOR023C	YOR019W	0.235222
YMR224C	YMR246W	0.220317	YOR028C	YOR013W	0.235222
YMR246W	YMR224C	0.220317	YOR037W	YPL259C	0.235222

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YOR054C	YLL055W	0.235222	YLR068W	YER060W-A	0.2681
YPL115C	YPL147W	0.235222	YML058C-A	YOL001W	0.2681
YPL125W	YPL159C	0.235222	YNL273W	YDR386W	0.2681
YPL147W	YPL115C	0.235222	YNL273W	YOR026W	0.2681
YPL159C	YPL125W	0.235222	YOL001W	YML058C-A	0.2681
YPL259C	YOR037W	0.235222	YOR026W	YNL273W	0.2681
YBR107C	YLR125W	0.246394	YPL072W	YBR195C	0.2681
YBR217W	YEL001C	0.246394	YPR135W	YCL061C	0.2681
YDR353W	YOR023C	0.246394	YBR173C	YPL260W	0.275556
YDR439W	YEL001C	0.246394	YDR125C	YOR054C	0.275556
YEL001C	YBR217W	0.246394	YDR353W	YPL001W	0.275556
YEL001C	YDR439W	0.246394	YDR388W	YNL294C	0.275556
YLR082C	YLR094C	0.246394	YEL053C	YPL001W	0.275556
YLR094C	YLR082C	0.246394	YGR061C	YHR111W	0.275556
YLR094C	YMR187C	0.246394	YHR111W	YGR061C	0.275556
YLR125W	YBR107C	0.246394	YHR114W	YLL039C	0.275556
YMR187C	YLR094C	0.246394	YLL039C	YHR114W	0.275556
YMR214W	YNR051C	0.246394	YMR198W	YMR263W	0.275556
YNR051C	YMR214W	0.246394	YMR263W	YMR198W	0.275556
YOR023C	YDR353W	0.246394	YNL294C	YDR388W	0.275556
YOR305W	YPL163C	0.246394	YOR054C	YDR125C	0.275556
YPL163C	YOR305W	0.246394	YPL001W	YDR353W	0.275556
YBR242W	YML058C-A	0.254157	YPL001W	YEL053C	0.275556
YBR244W	YMR246W	0.254157	YPL260W	YBR173C	0.275556
YER095W	YOR054C	0.254157	YEL037C	YLL015W	0.283991
YFR036W	YMR186W	0.254157	YLL015W	YEL037C	0.283991
YML058C-A	YBR242W	0.254157	YLL015W	YLL039C	0.283991
YMR144W	YMR176W	0.254157	YLL039C	YLL015W	0.283991
YMR176W	YMR144W	0.254157	YMR011W	YNL303W	0.283991
YMR186W	YFR036W	0.254157	YMR175W	YMR190C	0.283991
YMR246W	YBR244W	0.254157	YMR190C	YMR175W	0.283991
YOR054C	YER095W	0.254157	YNL303W	YMR011W	0.283991
YDR014W	YGL066W	0.258741	YOR010C	YOR034C	0.283991
YGL066W	YDR014W	0.258741	YOR034C	YOR010C	0.283991
YJL168C	YOR054C	0.258741	YBR009C	YBR107C	0.290497
YKL113C	YNL307C	0.258741	YBR107C	YBR009C	0.290497
YLR015W	YLR064W	0.258741	YBR194W	YJL047C	0.290497
YLR064W	YLR015W	0.258741	YDR359C	YPL152W	0.290497
YNL307C	YKL113C	0.258741	YER016W	YMR078C	0.290497
YOR054C	YJL168C	0.258741	YJL047C	YBR194W	0.290497
YDR334W	YNL330C	0.265589	YLL046C	YOL023W	0.290497
YMR244W	YNL275W	0.265589	YLR125W	YNL303W	0.290497
YNL275W	YMR244W	0.265589	YMR078C	YER016W	0.290497
YNL330C	YDR334W	0.265589	YNL303W	YLR125W	0.290497
YOR039W	YOR069W	0.265589	YOL023W	YLL046C	0.290497
YOR069W	YOR039W	0.265589	YPL152W	YDR359C	0.290497
YBR195C	YPL072W	0.2681	YBR231C	YDR359C	0.298319
YCL061C	YPR135W	0.2681	YDR080W	YOR068C	0.298319
YDL101C	YEL053C	0.2681	YDR358W	YMR284W	0.298319
YDR386W	YNL273W	0.2681	YDR359C	YBR231C	0.298319
YEL053C	YDL101C	0.2681	YDR393W	YMR275C	0.298319
YER060W-A	YLR068W	0.2681	YDR421W	YDR428C	0.298319

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YDR428C	YDR421W	0.298319	YPL145C	YMR176W	0.314777
YHR015W	YPR141C	0.298319	YPL145C	YMR246W	0.314777
YLL015W	YLL042C	0.298319	YDL074C	YKL113C	0.317383
YLL042C	YLL015W	0.298319	YDR079W	YPL270W	0.317383
YLL046C	YPL097W	0.298319	YDR117C	YJR043C	0.317383
YLR418C	YPL139C	0.298319	YDR359C	YPL155C	0.317383
YMR232W	YPL100W	0.298319	YER016W	YJR060W	0.317383
YMR275C	YDR393W	0.298319	YHR134W	YPL133C	0.317383
YMR284W	YDR358W	0.298319	YJR043C	YDR117C	0.317383
YNL297C	YOR067C	0.298319	YJR060W	YER016W	0.317383
YNL298W	YOR304W	0.298319	YKL113C	YDL074C	0.317383
YOR067C	YNL297C	0.298319	YLL020C	YOR037W	0.317383
YOR068C	YDR080W	0.298319	YLR068W	YOL011W	0.317383
YOR304W	YNL298W	0.298319	YLR418C	YMR198W	0.317383
YPL097W	YLL046C	0.298319	YML014W	YOR040W	0.317383
YPL100W	YMR232W	0.298319	YMR012W	YMR280C	0.317383
YPL139C	YLR418C	0.298319	YMR198W	YLR418C	0.317383
YPR141C	YHR015W	0.298319	YMR198W	YPL008W	0.317383
YDR126W	YPL232W	0.306653	YMR228W	YOR037W	0.317383
YGR188C	YOL076W	0.306653	YMR246W	YNL318C	0.317383
YLL024C	YOR014W	0.306653	YMR280C	YMR012W	0.317383
YLR021W	YNL298W	0.306653	YNL277W	YPL105C	0.317383
YMR228W	YMR264W	0.306653	YNL318C	YMR246W	0.317383
YMR264W	YMR228W	0.306653	YNR032C-A	YOR035C	0.317383
YNL298W	YLR021W	0.306653	YOL011W	YLR068W	0.317383
YOL076W	YGR188C	0.306653	YOR035C	YNR032C-A	0.317383
YOR014W	YLL024C	0.306653	YOR037W	YLL020C	0.317383
YPL232W	YDR126W	0.306653	YOR037W	YMR228W	0.317383
YAL039C	YOL055C	0.314777	YOR040W	YML014W	0.317383
YAL039C	YOR054C	0.314777	YPL008W	YMR198W	0.317383
YBR231C	YFR036W	0.314777	YPL105C	YNL277W	0.317383
YCL061C	YOR024W	0.314777	YPL125W	YPL145C	0.317383
YDR114C	YOR371C	0.314777	YPL133C	YHR134W	0.317383
YDR119W	YLR125W	0.314777	YPL145C	YPL125W	0.317383
YDR334W	YHR001W-A	0.314777	YPL155C	YDR359C	0.317383
YDR334W	YLR418C	0.314777	YPL155C	YPL269W	0.317383
YER060W-A	YOR054C	0.314777	YPL269W	YPL155C	0.317383
YFR036W	YBR231C	0.314777	YPL270W	YDR079W	0.317383
YHR001W-A	YDR334W	0.314777	YBR231C	YOR308C	0.319392
YLR125W	YDR119W	0.314777	YDL013W	YPL106C	0.319392
YLR418C	YDR334W	0.314777	YDR079W	YDR353W	0.319392
YMR176W	YPL145C	0.314777	YDR083W	YMR246W	0.319392
YMR246W	YPL145C	0.314777	YDR353W	YDR079W	0.319392
YOL011W	YOR014W	0.314777	YDR353W	YML007W	0.319392
YOL046C	YOR029W	0.314777	YDR359C	YDR385W	0.319392
YOL055C	YAL039C	0.314777	YDR385W	YDR359C	0.319392
YOR014W	YOL011W	0.314777	YGR061C	YPL116W	0.319392
YOR024W	YCL061C	0.314777	YLR011W	YLR016C	0.319392
YOR029W	YOL046C	0.314777	YLR015W	YPL125W	0.319392
YOR054C	YAL039C	0.314777	YLR016C	YLR011W	0.319392
YOR054C	YER060W-A	0.314777	YLR094C	YOR038C	0.319392
YOR371C	YDR114C	0.314777	YML007W	YDR353W	0.319392

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YMR223W	YOR025W	0.319392	YHR015W	YMR178W	0.329861
YMR246W	YDR083W	0.319392	YLR094C	YBR219C	0.329861
YNL298W	YNL330C	0.319392	YLR124W	YEL040W	0.329861
YNL330C	YNL298W	0.319392	YLR125W	YPL103C	0.329861
YOR025W	YMR223W	0.319392	YML013C-A	YNL297C	0.329861
YOR026W	YPR135W	0.319392	YML014W	YPL161C	0.329861
YOR038C	YLR094C	0.319392	YMR178W	YHR015W	0.329861
YOR308C	YBR231C	0.319392	YMR246W	YEL023C	0.329861
YPL097W	YPL101W	0.319392	YMR275C	YEL043W	0.329861
YPL101W	YPL097W	0.319392	YNL281W	YFR036W	0.329861
YPL106C	YDL013W	0.319392	YNL297C	YML013C-A	0.329861
YPL116W	YGR061C	0.319392	YNL298W	YOR360C	0.329861
YPL125W	YLR015W	0.319392	YOL046C	YOR040W	0.329861
YPR135W	YOR026W	0.319392	YOL053C-A	YPR141C	0.329861
YBR009C	YHR111W	0.319521	YOR035C	YAL039C	0.329861
YBR212W	YOR054C	0.319521	YOR040W	YOL046C	0.329861
YBR231C	YEL042W	0.319521	YOR054C	YGL194C	0.329861
YBR261C	YHR015W	0.319521	YOR067C	YPL103C	0.329861
YDL013W	YOR191W	0.319521	YOR360C	YNL298W	0.329861
YDR421W	YDR426C	0.319521	YPL100W	YDR334W	0.329861
YDR426C	YDR421W	0.319521	YPL103C	YLR125W	0.329861
YEL003W	YOR035C	0.319521	YPL103C	YOR067C	0.329861
YEL037C	YMR275C	0.319521	YPL147W	YBR173C	0.329861
YEL042W	YBR231C	0.319521	YPL161C	YBR200W	0.329861
YER016W	YML094W	0.319521	YPL161C	YML014W	0.329861
YGL066W	YPL101W	0.319521	YPR141C	YOL053C-A	0.329861
YHR015W	YBR261C	0.319521	YBR197C	YNL281W	0.330645
YHR031C	YPR120C	0.319521	YDL101C	YNL307C	0.330645
YHR111W	YBR009C	0.319521	YDR389W	YNL322C	0.330645
YHR191C	YMR048W	0.319521	YGL066W	YNL303W	0.330645
YLR089C	YMR179W	0.319521	YHR191C	YNL298W	0.330645
YML094W	YER016W	0.319521	YIL008W	YOL001W	0.330645
YMR048W	YHR191C	0.319521	YJL047C	YPL072W	0.330645
YMR179W	YLR089C	0.319521	YLL043W	YLR068W	0.330645
YMR275C	YEL037C	0.319521	YLR023C	YMR184W	0.330645
YNL297C	YOL054W	0.319521	YLR068W	YLL043W	0.330645
YOL054W	YNL297C	0.319521	YLR068W	YOL030W	0.330645
YOR035C	YEL003W	0.319521	YLR381W	YPR141C	0.330645
YOR054C	YBR212W	0.319521	YMR184W	YLR023C	0.330645
YOR191W	YDL013W	0.319521	YNL281W	YBR197C	0.330645
YPL101W	YGL066W	0.319521	YNL298W	YHR191C	0.330645
YPR120C	YHR031C	0.319521	YNL303W	YGL066W	0.330645
YAL039C	YOR035C	0.329861	YNL307C	YDL101C	0.330645
YBR173C	YPL147W	0.329861	YNL322C	YDR389W	0.330645
YBR200W	YPL161C	0.329861	YOL001W	YIL008W	0.330645
YBR219C	YLR094C	0.329861	YOL030W	YLR068W	0.330645
YDR334W	YPL100W	0.329861	YOL031C	YOR029W	0.330645
YEL023C	YMR246W	0.329861	YOL076W	YOR026W	0.330645
YEL040W	YLR124W	0.329861	YOR026W	YOL076W	0.330645
YEL043W	YMR275C	0.329861	YOR029W	YOL031C	0.330645
YFR036W	YNL281W	0.329861	YPL072W	YJL047C	0.330645
YGL194C	YOR054C	0.329861	YPL101W	YPL133C	0.330645

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YPL133C	YPL101W	0.330645	YOR019W	YNL314W	0.34736
YPR141C	YLR381W	0.330645	YOR040W	YEL029C	0.34736
YBR009C	YBR173C	0.338983	YOR291W	YNR051C	0.34736
YBR099C	YLR093C	0.338983	YOR334W	YML058C-A	0.34736
YBR173C	YBR009C	0.338983	YPL001W	YNL297C	0.34736
YDR334W	YOR308C	0.338983	YPL121C	YDR417C	0.34736
YDR353W	YOR360C	0.338983	YBR212W	YMR078C	0.353365
YEL043W	YMR265C	0.338983	YBR222C	YOR334W	0.353365
YGL163C	YOR054C	0.338983	YDL155W	YGR188C	0.353365
YGR188C	YML094W	0.338983	YDR014W	YEL043W	0.353365
YJL105W	YNL273W	0.338983	YDR108W	YNR051C	0.353365
YLL053C	YLR013W	0.338983	YDR123C	YOR038C	0.353365
YLR013W	YLL053C	0.338983	YDR130C	YML013C-A	0.353365
YLR093C	YBR099C	0.338983	YDR414C	YNL307C	0.353365
YLR093C	YNL311C	0.338983	YDR415C	YOR343C	0.353365
YML094W	YGR188C	0.338983	YEL008W	YPL125W	0.353365
YMR224C	YOR033C	0.338983	YEL043W	YDR014W	0.353365
YMR265C	YEL043W	0.338983	YEL043W	YPL100W	0.353365
YNL273W	YJL105W	0.338983	YGR188C	YDL155W	0.353365
YNL297C	YOR084W	0.338983	YLL052C	YOR054C	0.353365
YNL311C	YLR093C	0.338983	YML001W	YOL001W	0.353365
YOR033C	YMR224C	0.338983	YML013C-A	YDR130C	0.353365
YOR054C	YGL163C	0.338983	YMR011W	YOR009W	0.353365
YOR084W	YNL297C	0.338983	YMR078C	YBR212W	0.353365
YOR308C	YDR334W	0.338983	YMR157C	YMR228W	0.353365
YOR360C	YDR353W	0.338983	YMR198W	YOR026W	0.353365
YPL216W	YPL253C	0.338983	YMR228W	YMR157C	0.353365
YPL253C	YPL216W	0.338983	YMR275C	YMR278W	0.353365
YDR110W	YLR023C	0.34736	YMR278W	YMR275C	0.353365
YDR116C	YML009C	0.34736	YNL307C	YDR414C	0.353365
YDR383C	YGL060W	0.34736	YNR051C	YDR108W	0.353365
YDR417C	YPL121C	0.34736	YOL001W	YML001W	0.353365
YEL020C	YLR028C	0.34736	YOR009W	YMR011W	0.353365
YEL029C	YOR040W	0.34736	YOR026W	YMR198W	0.353365
YER095W	YNL294C	0.34736	YOR037W	YPL264C	0.353365
YGL060W	YDR383C	0.34736	YOR038C	YDR123C	0.353365
YGL194C	YMR283C	0.34736	YOR054C	YLL052C	0.353365
YJR060W	YLR418C	0.34736	YOR334W	YBR222C	0.353365
YLR015W	YMR012W	0.34736	YOR343C	YDR415C	0.353365
YLR023C	YDR110W	0.34736	YPL100W	YEL043W	0.353365
YLR028C	YEL020C	0.34736	YPL125W	YEL008W	0.353365
YLR418C	YJR060W	0.34736	YPL264C	YOR037W	0.353365
YML009C	YDR116C	0.34736	YBR073W	YLL015W	0.359253
YML058C-A	YOR334W	0.34736	YBR229C	YMR214W	0.359253
YMR012W	YLR015W	0.34736	YDL101C	YJR043C	0.359253
YMR273C	YNL302C	0.34736	YDR095C	YMR304W	0.359253
YMR283C	YGL194C	0.34736	YDR134C	YDR435C	0.359253
YNL294C	YER095W	0.34736	YDR353W	YOR038C	0.359253
YNL297C	YPL001W	0.34736	YDR359C	YHR191C	0.359253
YNL302C	YMR273C	0.34736	YDR435C	YDR134C	0.359253
YNL314W	YOR019W	0.34736	YEL001C	YOR305W	0.359253
YNR051C	YOR291W	0.34736	YEL001C	YPR141C	0.359253

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YEL008W	YLL039C	0.359253	YDR435C	YBR212W	0.372578
YHR015W	YMR160W	0.359253	YHR015W	YDR350C	0.372578
YHR191C	YDR359C	0.359253	YJR036C	YJR043C	0.372578
YJR043C	YDL101C	0.359253	YJR043C	YJR036C	0.372578
YLL015W	YBR073W	0.359253	YLR018C	YNL278W	0.372578
YLL039C	YEL008W	0.359253	YLR065C	YPL115C	0.372578
YLR021W	YOR054C	0.359253	YLR079W	YPR141C	0.372578
YLR053C	YLR056W	0.359253	YLR089C	YNL257C	0.372578
YLR056W	YLR053C	0.359253	YLR093C	YBR233W	0.372578
YLR119W	YNL021W	0.359253	YLR093C	YPL107W	0.372578
YML014W	YOR026W	0.359253	YLR125W	YMR233W	0.372578
YML094W	YOR014W	0.359253	YML058C-A	YOL052C	0.372578
YMR160W	YHR015W	0.359253	YMR233W	YLR125W	0.372578
YMR214W	YBR229C	0.359253	YMR297W	YDR097C	0.372578
YMR304W	YDR095C	0.359253	YNL257C	YLR089C	0.372578
YNL021W	YLR119W	0.359253	YNL273W	YOL067C	0.372578
YNL294C	YOR322C	0.359253	YNL278W	YLR018C	0.372578
YOR014W	YML094W	0.359253	YOL052C	YML058C-A	0.372578
YOR026W	YML014W	0.359253	YOL067C	YNL273W	0.372578
YOR038C	YDR353W	0.359253	YOR293W	YPL102C	0.372578
YOR054C	YLR021W	0.359253	YPL102C	YOR293W	0.372578
YOR305W	YEL001C	0.359253	YPL107W	YLR093C	0.372578
YOR322C	YNL294C	0.359253	YPL115C	YLR065C	0.372578
YPR141C	YEL001C	0.359253	YPR141C	YLR079W	0.372578
YBR173C	YEL001C	0.365649	YAL031C	YOR054C	0.389695
YCL016C	YMR198W	0.365649	YBR194W	YMR273C	0.389695
YDL013W	YML010W-A	0.365649	YCL016C	YNL298W	0.389695
YEL001C	YBR173C	0.365649	YDL134C	YDL188C	0.389695
YEL037C	YNL330C	0.365649	YDL188C	YDL134C	0.389695
YEL040W	YMR246W	0.365649	YDR116C	YOR010C	0.389695
YLL015W	YMR006C	0.365649	YDR383C	YMR163C	0.389695
YLR068W	YNL253W	0.365649	YDR388W	YML010W-A	0.389695
YLR093C	YML058C-A	0.365649	YGR188C	YOL023W	0.389695
YLR094C	YPL099C	0.365649	YHR015W	YMR148W	0.389695
YML010W-A	YDL013W	0.365649	YKR029C	YMR263W	0.389695
YML058C-A	YLR093C	0.365649	YLR068W	YMR274C	0.389695
YMR006C	YLL015W	0.365649	YML010W-A	YDR388W	0.389695
YMR179W	YMR223W	0.365649	YMR148W	YHR015W	0.389695
YMR198W	YCL016C	0.365649	YMR163C	YDR383C	0.389695
YMR223W	YMR179W	0.365649	YMR246W	YOR308C	0.389695
YMR246W	YEL040W	0.365649	YMR263W	YKR029C	0.389695
YNL253W	YLR068W	0.365649	YMR273C	YBR194W	0.389695
YNL330C	YEL037C	0.365649	YMR274C	YLR068W	0.389695
YOL014W	YOR023C	0.365649	YNL294C	YOL050C	0.389695
YOR023C	YOL014W	0.365649	YNL298W	YCL016C	0.389695
YPL099C	YLR094C	0.365649	YOL023W	YGR188C	0.389695
YBR173C	YDR358W	0.372578	YOL050C	YNL294C	0.389695
YBR212W	YDR435C	0.372578	YOR010C	YDR116C	0.389695
YBR233W	YLR093C	0.372578	YOR026W	YPL008W	0.389695
YDR097C	YMR297W	0.372578	YOR029W	YOR069W	0.389695
YDR350C	YHR015W	0.372578	YOR054C	YAL031C	0.389695
YDR358W	YBR173C	0.372578	YOR069W	YOR029W	0.389695

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YOR084W	YPL129W	0.389695	YDR353W	YBR200W	0.406421
YOR191W	YOR313C	0.389695	YEL014C	YLL043W	0.406421
YOR308C	YMR246W	0.389695	YEL053C	YML094W	0.406421
YOR313C	YOR191W	0.389695	YHR001W-A	YCR066W	0.406421
YOR382W	YPL216W	0.389695	YHR012W	YPL152W	0.406421
YPL008W	YOR026W	0.389695	YHR031C	YMR224C	0.406421
YPL129W	YOR084W	0.389695	YJR043C	YOR055W	0.406421
YPL216W	YOR382W	0.389695	YLL043W	YEL014C	0.406421
YBR201W	YPL096W	0.396893	YLR023C	YBR242W	0.406421
YBR261C	YMR304W	0.396893	YLR056W	YMR176W	0.406421
YDL074C	YGR188C	0.396893	YLR093C	YOR311C	0.406421
YDR200C	YLR093C	0.396893	YLR125W	YOR343C	0.406421
YEL018W	YEL024W	0.396893	YML094W	YDR334W	0.406421
YEL018W	YGL060W	0.396893	YML094W	YEL053C	0.406421
YEL024W	YEL018W	0.396893	YMR012W	YBR255W	0.406421
YGL060W	YEL018W	0.396893	YMR176W	YLR056W	0.406421
YGL163C	YML007W	0.396893	YMR198W	YDR200C	0.406421
YGR188C	YDL074C	0.396893	YMR198W	YMR224C	0.406421
YKL113C	YOR080W	0.396893	YMR224C	YHR031C	0.406421
YLR011W	YMR244W	0.396893	YMR224C	YMR198W	0.406421
YLR023C	YPL116W	0.396893	YNL269W	YOR013W	0.406421
YLR089C	YOR144C	0.396893	YNL335W	YOR069W	0.406421
YLR093C	YDR200C	0.396893	YOL004W	YDL074C	0.406421
YLR094C	YPL108W	0.396893	YOR003W	YDR116C	0.406421
YLR418C	YNL330C	0.396893	YOR013W	YNL269W	0.406421
YML007W	YGL163C	0.396893	YOR039W	YBL058W	0.406421
YMR224C	YPR135W	0.396893	YOR055W	YJR043C	0.406421
YMR244W	YLR011W	0.396893	YOR069W	YNL335W	0.406421
YMR246W	YPL108W	0.396893	YOR311C	YLR093C	0.406421
YMR304W	YBR261C	0.396893	YOR343C	YLR125W	0.406421
YNL330C	YLR418C	0.396893	YPL095C	YPL099C	0.406421
YOR029W	YOR033C	0.396893	YPL099C	YPL095C	0.406421
YOR033C	YOR029W	0.396893	YPL102C	YDL074C	0.406421
YOR080W	YKL113C	0.396893	YPL152W	YHR012W	0.406421
YOR144C	YLR089C	0.396893	YBR099C	YOL001W	0.415894
YPL096W	YBR201W	0.396893	YCL061C	YMR078C	0.415894
YPL097W	YPL139C	0.396893	YCR066W	YML006C	0.415894
YPL108W	YLR094C	0.396893	YDR014W	YKL113C	0.415894
YPL108W	YMR246W	0.396893	YDR097C	YPL097W	0.415894
YPL116W	YLR023C	0.396893	YDR110W	YDR117C	0.415894
YPL139C	YPL097W	0.396893	YDR117C	YDR110W	0.415894
YPR135W	YMR224C	0.396893	YDR386W	YML028W	0.415894
YBL058W	YOR039W	0.406421	YDR389W	YDR414C	0.415894
YBR200W	YDR353W	0.406421	YDR414C	YDR389W	0.415894
YBR242W	YLR023C	0.406421	YDR421W	YHR015W	0.415894
YBR255W	YMR012W	0.406421	YEL048C	YEL053C	0.415894
YCR066W	YHR001W-A	0.406421	YEL053C	YEL048C	0.415894
YDL074C	YOL004W	0.406421	YFR036W	YOR014W	0.415894
YDL074C	YPL102C	0.406421	YGR188C	YOR315W	0.415894
YDR116C	YOR003W	0.406421	YHR015W	YDR421W	0.415894
YDR200C	YMR198W	0.406421	YKL113C	YDR014W	0.415894
YDR334W	YML094W	0.406421	YKR029C	YPL097W	0.415894

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YLR065C	YOL001W	0.415894	YLR093C	YBR244W	0.420685
YML006C	YCR066W	0.415894	YLR125W	YBR195C	0.420685
YML028W	YDR386W	0.415894	YLR125W	YMR224C	0.420685
YML079W	YPL164C	0.415894	YML029W	YDR439W	0.420685
YML094W	YOL001W	0.415894	YML050W	YHR109W	0.420685
YMR078C	YCL061C	0.415894	YML079W	YPL145C	0.420685
YMR144W	YOR069W	0.415894	YMR010W	YOR005C	0.420685
YMR310C	YNL298W	0.415894	YMR198W	YDL155W	0.420685
YNL298W	YMR310C	0.415894	YMR224C	YLR125W	0.420685
YNL311C	YNL321W	0.415894	YMR282C	YLR068W	0.420685
YNL321W	YNL311C	0.415894	YNL265C	YPL001W	0.420685
YOL001W	YBR099C	0.415894	YNL303W	YEL003W	0.420685
YOL001W	YLR065C	0.415894	YOL004W	YJR043C	0.420685
YOL001W	YML094W	0.415894	YOL023W	YPL250C	0.420685
YOL013C	YOR039W	0.415894	YOL064C	YOR023C	0.420685
YOL055C	YPL119C	0.415894	YOR005C	YMR010W	0.420685
YOR014W	YFR036W	0.415894	YOR015W	YOR069W	0.420685
YOR014W	YOR359W	0.415894	YOR023C	YOL064C	0.420685
YOR039W	YOL013C	0.415894	YOR038C	YPL129W	0.420685
YOR069W	YMR144W	0.415894	YOR069W	YOR015W	0.420685
YOR315W	YGR188C	0.415894	YOR144C	YPL112C	0.420685
YOR359W	YOR014W	0.415894	YOR156C	YDL013W	0.420685
YPL097W	YDR097C	0.415894	YOR313C	YPL001W	0.420685
YPL097W	YKR029C	0.415894	YOR356W	YEL001C	0.420685
YPL119C	YOL055C	0.415894	YPL001W	YDR084C	0.420685
YPL164C	YML079W	0.415894	YPL001W	YNL265C	0.420685
YBR194W	YHR015W	0.420685	YPL001W	YOR313C	0.420685
YBR195C	YLR125W	0.420685	YPL101W	YPL239W	0.420685
YBR204C	YDR421W	0.420685	YPL112C	YLL016W	0.420685
YBR231C	YDR392W	0.420685	YPL112C	YOR144C	0.420685
YBR244W	YLR093C	0.420685	YPL129W	YOR038C	0.420685
YDL013W	YOR156C	0.420685	YPL130W	YLR023C	0.420685
YDL155W	YMR198W	0.420685	YPL145C	YML079W	0.420685
YDR084C	YPL001W	0.420685	YPL232W	YLR015W	0.420685
YDR130C	YPR068C	0.420685	YPL234C	YDR414C	0.420685
YDR392W	YBR231C	0.420685	YPL239W	YPL101W	0.420685
YDR414C	YPL234C	0.420685	YPL250C	YOL023W	0.420685
YDR421W	YBR204C	0.420685	YPR068C	YDR130C	0.420685
YDR439W	YML029W	0.420685	YBL058W	YHR015W	0.436959
YEL001C	YOR356W	0.420685	YBR194W	YOR360C	0.436959
YEL003W	YNL303W	0.420685	YCL016C	YFR036W	0.436959
YEL018W	YEL040W	0.420685	YDR051C	YDR435C	0.436959
YEL040W	YEL018W	0.420685	YDR096W	YMR228W	0.436959
YHR015W	YBR194W	0.420685	YDR334W	YDR359C	0.436959
YHR109W	YML050W	0.420685	YDR359C	YDR334W	0.436959
YIR004W	YLR015W	0.420685	YDR435C	YDR051C	0.436959
YJR043C	YOL004W	0.420685	YEL001C	YPL103C	0.436959
YLL016W	YPL112C	0.420685	YEL029C	YML058W	0.436959
YLR015W	YIR004W	0.420685	YEL037C	YMR283C	0.436959
YLR015W	YPL232W	0.420685	YEL039C	YML013C-A	0.436959
YLR023C	YPL130W	0.420685	YFR036W	YCL016C	0.436959
YLR068W	YMR282C	0.420685	YHR015W	YBL058W	0.436959

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YKR029C	YLR418C	0.436959	YMR198W	YMR144W	0.447716
YLL026W	YLL038C	0.436959	YMR246W	YLR089C	0.447716
YLL038C	YLL026W	0.436959	YNL269W	YDR357C	0.447716
YLL042C	YPR068C	0.436959	YNL288W	YLL039C	0.447716
YLL043W	YOL013C	0.436959	YOL001W	YOR038C	0.447716
YLR015W	YNL289W	0.436959	YOL004W	YGL058W	0.447716
YLR418C	YKR029C	0.436959	YOL011W	YOR054C	0.447716
YML013C-A	YEL039C	0.436959	YOL014W	YOR037W	0.447716
YML058W	YEL029C	0.436959	YOL046C	YOR024W	0.447716
YML094W	YPR141C	0.436959	YOR024W	YOL046C	0.447716
YMR187C	YPL113C	0.436959	YOR034C	YBR213W	0.447716
YMR228W	YDR096W	0.436959	YOR037W	YOL014W	0.447716
YMR283C	YEL037C	0.436959	YOR038C	YOL001W	0.447716
YNL273W	YOR083W	0.436959	YOR054C	YOL011W	0.447716
YNL289W	YLR015W	0.436959	YOR090C	YPL118W	0.447716
YOL013C	YLL043W	0.436959	YOR144C	YPL163C	0.447716
YOR083W	YNL273W	0.436959	YPL118W	YOR090C	0.447716
YOR311C	YOR347C	0.436959	YPL121C	YLR125W	0.447716
YOR347C	YOR311C	0.436959	YPL125W	YBR206W	0.447716
YOR360C	YBR194W	0.436959	YPL163C	YOR144C	0.447716
YPL102C	YPL234C	0.436959	YBR195C	YKL113C	0.456446
YPL103C	YEL001C	0.436959	YBR203W	YLR023C	0.456446
YPL113C	YMR187C	0.436959	YDR079W	YLL015W	0.456446
YPL234C	YPL102C	0.436959	YDR080W	YDR084C	0.456446
YPR068C	YLL042C	0.436959	YDR084C	YDR080W	0.456446
YPR141C	YML094W	0.436959	YDR093W	YML013W	0.456446
YBR009C	YDR083W	0.447716	YDR096W	YLR020C	0.456446
YBR194W	YLR113W	0.447716	YDR096W	YPL102C	0.456446
YBR200W	YDR414C	0.447716	YDR107C	YOR038C	0.456446
YBR206W	YPL125W	0.447716	YDR117C	YLR011W	0.456446
YBR213W	YOR034C	0.447716	YDR353W	YML008C	0.456446
YCL016C	YJR043C	0.447716	YDR389W	YPL234C	0.456446
YCL061C	YKL113C	0.447716	YDR439W	YML013C-A	0.456446
YDR072C	YDR080W	0.447716	YEL031W	YIR004W	0.456446
YDR072C	YLR057W	0.447716	YER016W	YLR107W	0.456446
YDR080W	YDR072C	0.447716	YER059W	YLR068W	0.456446
YDR083W	YBR009C	0.447716	YGL060W	YMR198W	0.456446
YDR123C	YFR036W	0.447716	YHR015W	YHR034C	0.456446
YDR357C	YNL269W	0.447716	YHR034C	YHR015W	0.456446
YDR414C	YBR200W	0.447716	YHR111W	YOL004W	0.456446
YFR036W	YDR123C	0.447716	YIR004W	YEL031W	0.456446
YGL058W	YOL004W	0.447716	YKL113C	YBR195C	0.456446
YJL047C	YKL113C	0.447716	YLL015W	YDR079W	0.456446
YJR043C	YCL016C	0.447716	YLL015W	YPL136W	0.456446
YKL113C	YCL061C	0.447716	YLR011W	YDR117C	0.456446
YKL113C	YJL047C	0.447716	YLR015W	YPL108W	0.456446
YLL039C	YNL288W	0.447716	YLR020C	YDR096W	0.456446
YLR057W	YDR072C	0.447716	YLR023C	YBR203W	0.456446
YLR089C	YMR246W	0.447716	YLR068W	YER059W	0.456446
YLR113W	YBR194W	0.447716	YLR107W	YER016W	0.456446
YLR125W	YPL121C	0.447716	YLR418C	YML010W-A	0.456446
YMR144W	YMR198W	0.447716	YML008C	YDR353W	0.456446

88

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YML010W-A	YLR418C	0.456446	YJR036C	YMR246W	0.46882
YML013C-A	YDR439W	0.456446	YJR060W	YDL074C	0.46882
YML013W	YDR093W	0.456446	YKL113C	YEL029C	0.46882
YMR198W	YGL060W	0.456446	YLL047W	YLR015W	0.46882
YMR253C	YOR311C	0.456446	YLR015W	YBR219C	0.46882
YNR032C-A	YOR079C	0.456446	YLR015W	YBR261C	0.46882
YOL004W	YHR111W	0.456446	YLR015W	YDR386W	0.46882
YOR014W	YPL253C	0.456446	YLR015W	YLL047W	0.46882
YOR038C	YDR107C	0.456446	YLR023C	YOR191W	0.46882
YOR079C	YNR032C-A	0.456446	YLR059C	YOR054C	0.46882
YOR311C	YMR253C	0.456446	YLR125W	YPL155C	0.46882
YOR365C	YOR375C	0.456446	YLR418C	YOR065W	0.46882
YOR375C	YOR365C	0.456446	YML028W	YGL163C	0.46882
YPL092W	YPL100W	0.456446	YML079W	YDR439W	0.46882
YPL100W	YPL092W	0.456446	YML084W	YPL110C	0.46882
YPL102C	YDR096W	0.456446	YMR148W	YNL297C	0.46882
YPL108W	YLR015W	0.456446	YMR179W	YEL053C	0.46882
YPL133C	YPL147W	0.456446	YMR246W	YJR036C	0.46882
YPL136W	YLL015W	0.456446	YMR265C	YMR278W	0.46882
YPL147W	YPL133C	0.456446	YMR273C	YDR142C	0.46882
YPL234C	YDR389W	0.456446	YMR278W	YMR265C	0.46882
YPL253C	YOR014W	0.456446	YNL021W	YDR388W	0.46882
YBR194W	YNL288W	0.46882	YNL288W	YBR194W	0.46882
YBR200W	YDR435C	0.46882	YNL288W	YPL239W	0.46882
YBR219C	YLR015W	0.46882	YNL297C	YMR148W	0.46882
YBR261C	YLR015W	0.46882	YOL004W	YOR065W	0.46882
YBR264C	YOR028C	0.46882	YOL036W	YOR026W	0.46882
YCL016C	YCL061C	0.46882	YOR023C	YPL086C	0.46882
YCL016C	YGL058W	0.46882	YOR026W	YOL036W	0.46882
YCL061C	YCL016C	0.46882	YOR028C	YBR264C	0.46882
YDL074C	YDR334W	0.46882	YOR054C	YLR059C	0.46882
YDL074C	YJR060W	0.46882	YOR065W	YLR418C	0.46882
YDR080W	YDR097C	0.46882	YOR065W	YOL004W	0.46882
YDR097C	YDR080W	0.46882	YOR191W	YLR023C	0.46882
YDR142C	YMR273C	0.46882	YOR322C	YEL020C	0.46882
YDR334W	YDL074C	0.46882	YPL086C	YOR023C	0.46882
YDR386W	YLR015W	0.46882	YPL110C	YML084W	0.46882
YDR388W	YNL021W	0.46882	YPL116W	YDR414C	0.46882
YDR414C	YPL116W	0.46882	YPL155C	YLR125W	0.46882
YDR435C	YBR200W	0.46882	YPL239W	YNL288W	0.46882
YDR439W	YML079W	0.46882	YAL036C	YMR246W	0.479744
YEL001C	YGR061C	0.46882	YBR194W	YOR380W	0.479744
YEL020C	YOR322C	0.46882	YBR195C	YLR107W	0.479744
YEL029C	YKL113C	0.46882	YBR219C	YOR054C	0.479744
YEL042W	YHR012W	0.46882	YDR079W	YOR010C	0.479744
YEL049W	YHR134W	0.46882	YDR093W	YOR003W	0.479744
YEL053C	YMR179W	0.46882	YDR094W	YML058C-A	0.479744
YGL058W	YCL016C	0.46882	YDR096W	YNL291C	0.479744
YGL163C	YML028W	0.46882	YDR116C	YOR007C	0.479744
YGR061C	YEL001C	0.46882	YDR171W	YNL288W	0.479744
YHR012W	YEL042W	0.46882	YDR200C	YML013C-A	0.479744
YHR134W	YEL049W	0.46882	YDR259C	YLL039C	0.479744

ORF1	ORF2	FDR	ORF1	ORF2	FDR
YDR359C	YOR035C	0.479744	YOR156C	YOR144C	0.479744
YDR388W	YEL031W	0.479744	YOR304W	YLL021W	0.479744
YEL017W	YEL031W	0.479744	YOR311C	YMR285C	0.479744
YEL031W	YDR388W	0.479744	YOR322C	YNR051C	0.479744
YEL031W	YEL017W	0.479744	YOR322C	YOR346W	0.479744
YEL043W	YLL046C	0.479744	YOR334W	YOL045W	0.479744
YJR036C	YLR053C	0.479744	YOR346W	YOR322C	0.479744
YLL021W	YLR015W	0.479744	YOR350C	YML079W	0.479744
YLL021W	YOR304W	0.479744	YOR380W	YBR194W	0.479744
YLL039C	YDR259C	0.479744	YPL072W	YMR246W	0.479744
YLL046C	YEL043W	0.479744	YPL072W	YPL129W	0.479744
YLR015W	YLL021W	0.479744	YPL102C	YOR038C	0.479744
YLR021W	YMR198W	0.479744	YPL111W	YMR246W	0.479744
YLR053C	YJR036C	0.479744	YPL129W	YPL072W	0.479744
YLR077W	YOR054C	0.479744	YPL254W	YNR032C-A	0.479744
YLR102C	YNL246W	0.479744	YPR135W	YNL298W	0.479744
YLR107W	YBR195C	0.479744	YAL039C	YMR179W	0.488181
YLR125W	YMR251W-A	0.479744	YBL058W	YMR145C	0.488181
YML013C-A	YDR200C	0.479744	YBR197C	YPL216W	0.488181
YML013W	YNL289W	0.479744	YBR200W	YNR051C	0.488181
YML058C-A	YDR094W	0.479744	YBR227C	YOR029W	0.488181
YML079W	YOR350C	0.479744	YBR231C	YKR029C	0.488181
YMR012W	YNR032C-A	0.479744	YCL061C	YMR048W	0.488181
YMR198W	YLR021W	0.479744	YDL013W	YDR359C	0.488181
YMR246W	YAL036C	0.479744	YDR014W	YMR198W	0.488181
YMR246W	YNL335W	0.479744	YDR094W	YDR111C	0.488181
YMR246W	YPL072W	0.479744	YDR094W	YMR176W	0.488181
YMR246W	YPL111W	0.479744	YDR094W	YOL064C	0.488181
YMR251W-A	YLR125W	0.479744	YDR096W	YML014W	0.488181
YMR285C	YOR311C	0.479744	YDR110W	YML084W	0.488181
YNL246W	YLR102C	0.479744	YDR111C	YDR094W	0.488181
YNL246W	YNL277W	0.479744	YDR358W	YEL053C	0.488181
YNL277W	YNL246W	0.479744	YDR359C	YDL013W	0.488181
YNL288W	YDR171W	0.479744	YDR395W	YDR426C	0.488181
YNL289W	YML013W	0.479744	YDR406W	YPL072W	0.488181
YNL291C	YDR096W	0.479744	YDR426C	YDR395W	0.488181
YNL298W	YPR135W	0.479744	YEL008W	YOR039W	0.488181
YNL335W	YMR246W	0.479744	YEL053C	YDR358W	0.488181
YNR032C-A	YMR012W	0.479744	YHR034C	YOR005C	0.488181
YNR032C-A	YPL254W	0.479744	YKR029C	YBR231C	0.488181
YNR051C	YOR322C	0.479744	YLR023C	YMR278W	0.488181
YOL045W	YOR334W	0.479744	YLR023C	YPL156C	0.488181
YOL053C-A	YOR002W	0.479744	YLR119W	YPL125W	0.488181
YOR002W	YOL053C-A	0.479744	YML013C-A	YPL103C	0.488181
YOR003W	YDR093W	0.479744	YML013C-A	YPL120W	0.488181
YOR007C	YDR116C	0.479744	YML014W	YDR096W	0.488181
YOR010C	YDR079W	0.479744	YML084W	YDR110W	0.488181
YOR035C	YDR359C	0.479744	YMR008C	YPL140C	0.488181
YOR038C	YPL102C	0.479744	YMR048W	YCL061C	0.488181
YOR054C	YBR219C	0.479744	YMR145C	YBL058W	0.488181
YOR054C	YLR077W	0.479744	YMR176W	YDR094W	0.488181
YOR144C	YOR156C	0.479744	YMR179W	YAL039C	0.488181

ORF1	ORF2	FDR
YMR198W	YDR014W	0.488181
YMR233W	YOR007C	0.488181
YMR256C	YOR308C	0.488181
YMR278W	YLR023C	0.488181
YMR280C	YMR285C	0.488181
YMR285C	YMR280C	0.488181
YNL281W	YNL294C	0.488181
YNL294C	YNL281W	0.488181
YNR051C	YBR200W	0.488181
YOL014W	YOR024W	0.488181
YOL030W	YPL118W	0.488181
YOL064C	YDR094W	0.488181
YOR005C	YHR034C	0.488181
YOR007C	YMR233W	0.488181
YOR011W	YPL118W	0.488181
YOR023C	YPR141C	0.488181
YOR024W	YOL014W	0.488181
YOR029W	YBR227C	0.488181
YOR039W	YEL008W	0.488181
YOR081C	YPL254W	0.488181
YOR308C	YMR256C	0.488181
YPL072W	YDR406W	0.488181
YPL103C	YML013C-A	0.488181
YPL118W	YOL030W	0.488181
YPL118W	YOR011W	0.488181
YPL120W	YML013C-A	0.488181
YPL125W	YLR119W	0.488181
YPL140C	YMR008C	0.488181
YPL156C	YLR023C	0.488181
YPL216W	YBR197C	0.488181
YPL254W	YOR081C	0.488181
YPR141C	YOR023C	0.488181

Table 3. Predicted synthetic rescue pairs by estimated FDR.

ORF1	ORF2	FDR
YOR023C	YOR034C	0.1
YOR034C	YOR023C	0.1
YLR023C	YMR255W	0.125
YMR255W	YLR023C	0.125
YOR023C	YOR035C	0.166666667
YOR035C	YOR023C	0.166666667
YMR242C	YNL302C	0.388888889
YNL302C	YMR242C	0.388888889
YDR114C	YPL097W	0.4
YPL097W	YDR114C	0.4
YGR188C	YHR110W	0.416666667
YHR110W	YGR188C	0.416666667
YPL106C	YPL239W	0.416666667
YPL239W	YPL106C	0.416666667
YML013C-A	YOR039W	0.428571429
YML020W	YML024W	0.428571429
YML024W	YML020W	0.428571429
YOR039W	YML013C-A	0.428571429
YMR242C	YPL239W	0.4375
YPL239W	YMR242C	0.4375
YBR194W	YNL021W	0.461538462
YNL021W	YBR194W	0.461538462

Supplementary Table 1. TAGs annotated as failed.

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YBL006C	Dn	AGACTACTTGAACGATCCTC	YCR057C	Dn	CACTAGCGATAAGTTCTGTC
YBL013W	Dn	AAGACCGACTAACTGATCTC	YCR090C	Dn	GATGTTTCACCATAGTCCTC
YBL014C	Dn	ATAGTACGACCAGGACTCTC	YCR091W	Dn	TCACGTTACTGAATGTCCTC
YBL021C	Dn	AGGACATCTATAACTCCGTC	YDL001W	Dn	CATACGCGGTAAGGATATAG
YBL026W	Dn	AGCAGACTCCTTAATTCGTC	YDL042C	Dn	CGCTATTGCTTAATCGTAGG
YBL029W	Dn	AATCATCGGTGGCAACGGTC	YDL044C	Dn	CAGATACTATTAAGTGCCGG
YBL046W	Dn	ACCACTTATGTGAAGCGTTC	YDL045C	Dn	CTTACTTGTAGCATAGAGGG
YBL055C	Dn	AGCGCACCCATTCGGATAAG	YDL049C	Dn	CATCATTTGTGGTAATCGGG
YBL072C	Dn	ACCTGTAGAGTAAGTGTCAG	YDL069C	Dn	CCATTGTGTTAAGGGTCATG
YBL075C	Dn	AGGCTATACCCATTCTTCAG	YDL075W	Dn	CACTAGGTATGGTAAGACTG
YBL090W	Dn	ACCGCGCTCTAATACTTGAG	YDL077C	Dn	CGCGGCAGTTATTTAATCTG
YBL096C	Dn	AAGACACCAGCCACTGTAGG	YDL092W	Dn	CACCTGTGTTGTAAGGATTG
YBR018C	Dn	ATATGACCTACGCCATCGGG	YDL094C	Dn	CATACTACGTGGGACAGTTG
YBR030W	Dn	ATATACCTCTCAGGCGGTGG	YDL132W	Dn	CCTCTGTTCGAGTGGAGATT
YBR032W	Dn	AATGGAACCGGCCTGACATG	YDL163W	Dn	CGGAGCGTTATGATACTGTT
YBR034C	Dn	ATTAGGGTCCACTGCCCATG	YDL166C	Dn	CGAGCTACGCTACTATTGTT
YBR038W	Dn	ACCTCCTAGAGAGTGGTATG	YDL192W	Dn	GCCTATACCCAACTAATGGA
YBR040W	Dn	ATCCGACGTTGCAGGAACTG	YDL196W	Dn	GTGAACAATAACGGCCTTGA
YBR041W	Dn	AGTAGGATAGCCTCCCACTG	YDL198C	Dn	GCCCGAACAGTTGTGAAATA
YBR049C	Dn	ATGATGGGTTCCACTCCCTG	YDL199C	Dn	GGTCGGACACTACACCAATA
YBR070C	Dn	ATACCACCCTCAGTAGTGTG	YDL206W	Dn	GGAGAATCTACCGCAACCTA
YBR075W	Dn	CGACTCGACTTAGATATGTG	YDL213C	Dn	GAATTGGGTACAAGCCTCTA
YBR085W	Dn	ATGGTCCGACTGGATCACCT	YDL216C	Dn	GCGCCTATTACACAAACGTA
YBR101C	Dn	AGGTGGTTCTCACTCATCCT	YDL223C	Dn	GCTACCTCAAGACGTACTTA
YBR105C	Dn	ATTCCGAGTGCCCTGAAGCT	YDL229W	Dn	GCCCAGTGTCAATTTCGTTA
YBR122C	Dn	ACTATTTGAGTCAGTGGGCT	YDR005C	Dn	GATTCTCAATAAGCCCGAAC
YBR123C	Dn	AACTGGATTGCATTTGGGCT	YDR006C	Dn	GCGTGTGAATAATACGGAAC
YBR126C	Dn	ACTTGAGGCGTAGACCTGCT	YDR027C	Dn	GACTAACCCACATTGGTGAC
YBR133C	Dn	AAACCTCGCGTCATTGATCT	YDR033W	Dn	GAAGCATACCCACGTTATAC
YBR152W	Dn	AGTTCACCATGCCTAGTTCT	YDR034C	Dn	GAAGGATCTCGAACACCTAC
YBR192W	Dn	ACGCTACCGCTGATCTATTT	YDR037W	Dn	GATAGTGAACCACCTTCTAC
YBR200W	Dn	ACGACTCTGGGATGATGTTT	YDR045C	Dn	GCGCTCGACTAAGAGAAACC
YBR202W	Dn	CCTCGCTCCAAATTGAGAAA	YDR046C	Dn	GCGCATAATGTTTACCAACC
YBR214W	Dn	CCTATCCGCAAAGTGGAGAA	YDR052C	Dn	GGCTTATACTAATCTCCCAG
YBR216C	Dn	CGCCCGTGCTAACAATAGAA	YDR054C	Dn	GAACCTAGCGTAATACTGAG
YBR262C	Dn	CCGTACTGCAAAGAAGGGCA	YDR055W	Dn	GATCCACGGACCTCGATTAG
YBR271W	Dn	CCTAGTAATCACGATCTTGC	YDR086C	Dn	GAAGCAAACACCATCCTTGG
YBR273C	Dn	GGACGATCATCCAGCACTAG	YDR089W	Dn	GTGTCACCTACCGTACTATG
YBR274W	Dn	GGCTCATCTCTAATCACTAG	YDR090C	Dn	GACGGTACTCCATTTGTATG
YBR285W	Dn	TGATGCAGAGTTCTCGGAAT	YDR091C	Dn	GAGGGTTCTCCACTTCACTG
YBR299W	Dn	ACCCTTCTGAGACGGTAGTG	YDR092W	Dn	GGTCGATATAGACGTTACTG
YCL031C	Dn	CCGATTTCGCAACATGGTTA	YDR111C	Dn	GCTGACTACTTACCGCTAAT
YCL052C	Dn	CCTTAGCGAGAATAGGAGAC	YDR112W	Dn	GTACGCGCAGTATCGGTAAT
YCL074W	Dn	CAGGGAATCGAATCTACTAC	YDR129C	Dn	GATTTGGCTCCAGGCCCTAT
YCR003W	Dn	CAGTCGGTAGAACTATGTAC	YDR130C	Dn	GTTCAGTGACGACTCCCTAT
YCR021C	Dn	CGCACCTAGATAAGATTTCC	YDR138W	Dn	GACCTAGCTTTACAGTCACT
YCR024C-A	Dn	CAATGCGACCCAAGGTGTGA	YDR145W	Dn	GGTATCGCTATATCCCTACT
YCR027C	Dn	GCGCGGCTGCAATTATATTA	YDR154C	Dn	TGCGCCGTACCAAGAAACGA

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YDR157W	Dn	TCGCGGTTACAAGATAAGGA	YFR027W	Dn	CGTAGCATAGCACTAGATAG
YDR160W	Dn	TCGCCAGGTACACAAAGGGA	YFR028C	Dn	GCACTACTACTACACGATAG
YDR175C	Dn	TGGGAGACCTAACACCATAC	YFR037C	Dn	TGACAGTCCACATAGTTCTC
YDR184C	Dn	TACGCAAATCAAGGTTAGCC	YFR038W	Dn	TCAGTGGACACAGCGTTCTC
YDR204W	Dn	TTCACGAGACTAAGGGCCTC	YFR039C	Dn	GAGCAGTCCCTTAATTTCTC
YDR210W	Dn	TAAGGACTGATAAGCCGGTC	YFR040W	Dn	CATGACAGATTGGACAAGTC
YDR217C	Dn	TGGTTACCCATCTTACACAG	YGL004C	Dn	CCGAGACCTTATCAGGAAAT
YDR227W	Dn	TACACTCGACGACACGGTAG	YGL005C	Dn	CGACCCGACTTCTATGAAAT
YDR290W	Dn	TCTAAGAACCGCCTTGCAGT	YGL006W	Dn	CACCCGGAGATTGAGTAAAT
YDR313C	Dn	TTAGCCTTCTCATGGCCGGT	YGL009C	Dn	CCTAGATGACGTTTGGGAAT
YDR335W	Dn	TCAGTTTGGGCAGTCCCTGT	YGL018C	Dn	CCGGTCTCTTAATATGGACT
YDR356W	Dn	ATATCACCGGGCACGGGTAT	YGL034C	Dn	CTTATCTGAGTTAGGACGGT
YDR367W	Dn	ATTAGTGGGTCCATCGCCCT	YGL044C	Dn	CACGGGACTTTGTGAGAATT
YDR394W	Dn	AAGTCTTCGGCATACCTTCT	YGL066W	Dn	CAACCACTCCGGTATATGTT
YDR398W	Dn	AAGTCAGTTGCACCCTTTCT	YGL093W	Dn	GCCGAAACCCAATTTGTCAA
YDR402C	Dn	AGTTGCTATACCTCCCTAGT	YGL097W	Dn	GGCCTACTCAAACCTTAGAA
YDR427W	Dn	CCCTATGGTTGTGAATGATG	YGL097W	Dn	GGCCTACTCAAACCTTAGAA
YDR433W	Dn	AATACTCTCTGACGGGAGGT	YGL099W	Dn	GGCGTTCGTACATCCCTTAT
YDR441C	Dn	CCCGCGTAGAAATTACTGAA	YGL101W	Dn	GCGGGCACCTTATTATACCT
YDR442W	Dn	CCGAACGGACAATTTCTGAA	YGL104C	Dn	GGGACGGACTTCATTCTCCT
YDR448W	Dn	CTGCGGGCGAAAGTTAATAA	YGL129C	Dn	GCTACGGTTAGACTTGCGTT
YDR456W	Dn	CCCGACGAGAAATGTTGTAA	YGL136C	Dn	GAAGGATCGTTCGGCACTTT
YDR465C	Dn	CCGTACCGGAAATCTATACA	YGL167C	Dn	TAAGTGGGACTAAGGTCTTC
YDR499W	Dn	CTTTAAGGGCAATACAGGGA	YGL179C	Dn	TATTAGGTACGCCCACCAGG
YDR501W	Dn	CACGTCGTTAAAGACCTGGA	YGL198W	Dn	AACCACTCTTTAACGTGAGG
YDR507C	Dn	CCCGGAAGGCAATTACACTA	YGL200C	Dn	AGTTCACGACGATTACCCGG
YDR516C	Dn	CCGGCAGGTCAATTAAGTTA	YGL210W	Dn	AGGGTAGTCCACTTTCCCAT
YDR518W	Dn	CCACGCGAAGAACGGATTTA	YGL216W	Dn	AATGGTAGGTGACCCTCCCT
YEL005C	Dn	CAGCCGGTTAGATATGATTG	YGL220W	Dn	AATTGGAGGGCACCCGTTCT
YEL006W	Dn	GCAGCTACTCGCACTGATTG	YGL226C-A	Dn	ACGGTGACTTACCCTCAGGT
YEL009C	Dn	CCCGGCATCTGCTATAATAT	YGL230C	Dn	ATCCCGGTTGCAGTTTCGGT
YEL017W	Dn	GGCGTCACTAAATCCAATCA	YGL245W	Dn	ACCGTGTACCTACTCTAGTT
YEL028W	Dn	CGCCCGACATGATGAAAGTA	YGR002C	Dn	ACCTGGCGTTCACGTAGTTT
YEL038W	Dn	AAGTCAGTCGAATGCCCTGC	YGR012W	Dn	CCGCCGTTGAAAGAACCTAA
YEL061C	Dn	GGATTGCATCTATCGTCAGT	YGR015C	Dn	CTCCCTCCTAAAGAAGGTAA
YER025W	Dn	TGCCTGTCAACAATCAGCAA	YGR020C	Dn	CCGGCGGATAAACGTAACCA
YER030W	Dn	GCTGGATATAAATCAGGCGA	YGR024C	Dn	CCAAGTACCGTAAACGTCGA
YER031C	Dn	GGCTCCACTAAATAGACGCA	YGR036C	Dn	AAACCTACGTTGGATAAGGG
YER038C	Dn	TGCATTAACACATACGGCAC	YGR037C	Dn	AACTAGGTTTAAGGTCCTGG
YER049W	Dn	CATGACTACTAAGGCGTATC	YGR047C	Dn	CCCGGTACGTTAATTTGTAG
YER067W	Dn	AAGCCGATGGCATGTCAGAT	YGR062C	Dn	GTTGGTCATCCACAAACACA
YER077C	Dn	GCCCGAGGCAAATTCAGCAA	YGR068C	Dn	CCCGCCTTTAGAGAAATACA
YER092W	Dn	ATATCGGCAGAAGCTGGCAC	YGR079W	Dn	GATAAGAGCACAGCTTCCCA
YER093C	Dn	GGATCACCAGAATCATGCAC	YGR085C	Dn	TCCGAATTAACAATCCGGCA
YER147C	Dn	TGACACCATATACAGGAGTC	YGR086C	Dn	GGGACTTCACAACCAATGCA
YER162C	Dn	GGCACGTTGTCTAACTACGG	YGR090W	Dn	GCGCATAAAGACACGTTGCA
YFL046W	Dn	TCCACGGGATACAGTCTGAG	YGR097W	Dn	GGCCACCGACAATGAAGTCA
YFR002W	Dn	GGCCCTTTAAGATCATTGAG	YGR110W	Dn	GGCTCTCCGAACATACAAGA
YFR004W	Dn	CAGCATCTACCATTCTTGAG	YGR115C	Dn	GGGCCGCACTTACAAACAGA
YFR010W	Dn	ACAGACACCTTGCTCAATAG	YGR117C	Dn	TCGCTTCAACCAAGGACAGA
YFR024C-A	Dn	AGCCAGCTAGTGTAAGATAG	YGR122C-A	Dn	AGAGCATCTGTCCTAACGTG

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YGR123C	Dn	CGCCCACAACAATAGTTTGA	YIL134W	Dn	CCTCGATAGCAATGACCATA
YGR133W	Dn	CGTCCGCCGAAGATTAAATA	YIL147C	Dn	CAAGATAGGCTAACAGTCGC
YGR152C	Dn	GTGCCCACACTCAAGAGATA	YIL148W	Dn	GCATTCACGATAACGGCATC
YGR155W	Dn	GAGTCTCACCGACACTGATG	YIL150C	Dn	ATAGCCGCTGAAGTGGCATC
YGR178C	Dn	GCAAGACTTAGCAACTCCTA	YIL156W	Dn	GCCTGATGTATAAGCAGTTC
YGR179C	Dn	AAGAACGCTACAGCTTCCTA	YIR001C	Dn	ATGGATACGTGCTGTAAGCT
YGR195W	Dn	GCTCATCTCCAAGGTATCTA	YIR010W	Dn	GATACGCACCAAGTCTCAGA
YGR198W	Dn	GTCAACCGTAACGTACTCTA	YJL005W	Dn	CTGCTGTGAAGACTGTTTAG
YGR200C	Dn	ATACGTGGACAAGCGGTCTA	YJL008C	Dn	CGAGAGTGACTAACTGCTTC
YGR217W	Dn	CCCGCTATCAGAAGAACGTA	YJL016W	Dn	AATCATGCTGAACTGCCATC
YGR226C	Dn	TACAGGCGGTAATTGCCGTC	YJL034W	Dn	GCGATTAGATGCAGTCTGAT
YGR228W	Dn	AACAGAGCTTTAACAGCGTC	YJL069C	Dn	ACACGATGACAATGCCTGGA
YGR230W	Dn	TGGACTCTCATAACGGCGTC	YJL081C	Dn	GGACTATCGTCATGCTGTCT
YGR244C	Dn	ACGTAAGCTAGAGTAAGGTC	YJL116C	Dn	CCCATGATCTCAATTAGGAC
YGR248W	Dn	AACCATTCCTGAAGACGGTC	YJL117W	Dn	GATCCGAGAACATCACTCTA
YGR261C	Dn	CTACAGGGTACAGTATCCCG	YJL122W	Dn	CTCGACAGCAGAACACTGGA
YHL012W	Dn	ACTGAGTGTGAATCATGGTC	YJL126W	Dn	AGATTCCTGAAAGCCCTGCA
YHL039W	Dn	GCTCCGAACCAATAATGTCA	YJL127C	Dn	GCCTATGACAAAGACCTGCA
YHR005C	Dn	TCCCTGATGGGTGATAAGAT	YJL136C	Dn	CCTGTCCAGAAAGCCATGAA
YHR008C	Dn	TAGACTGGCGCAGGTATTAT	YJL140W	Dn	TGGCAGGTCTCATGCTCTCT
YHR036W	Dn	CATGTATCAGAACGCATCAC	YJL163C	Dn	ACCGATCTGGAATGATATGC
YHR040W	Dn	TCACGGAATCGAGATGAAGC	YJL165C	Dn	CGCTATACCAGCAAATATGC
YHR041C	Dn	TAGTAGTGCTCAGCGTCCCT	YJL167W	Dn	GCCTTATATGAACTCTCAGC
YHR050W	Dn	CGGCGCGATTATGATGAAAT	YJL170C	Dn	CATGGTAATGAAGCATCAGC
YHR057C	Dn	CAATGATGCGTCAGTTAGGT	YJL172W	Dn	TAGGGATCGCTACCCATCTG
YHR061C	Dn	CCTCAGTAGAAAGCTGGACA	YJL189W	Dn	CCACACTGCAAATCTGGGAA
YHR064C	Dn	CCCATGCAGAAAGGCTGACA	YJL191W	Dn	AGCGACTGATTTCCATCTCT
YHR085W	Dn	ACTGCATGTGAAGCCATTGC	YJL194W	Dn	CGAGACTGTCGCATGATGAT
YHR088W	Dn	ATAGTGCAGCCAGCCCATAG	YJL196C	Dn	TCTTGCAGGCGCTAGATGAT
YHR091C	Dn	GGCCAGACTCTACATCATAG	YJL218W	Dn	AGATATTGACCACCATCAGC
YHR097C	Dn	ATTGTGCGTACACTGCCCTG	YJR007W	Dn	ACTTCAGTAGAGGTGAGCAT
YHR115C	Dn	GCGCAAGACACACGTTGACA	YJR008W	Dn	TAGCTCAGGCTCAGACGCAT
YHR150W	Dn	TCTCGTAGAGGACTGTAGAT	YJR010C-A	Dn	CATACTCTAGTGCATTGCCT
YHR152W	Dn	TTTGTGCGACGCTCACCGAT	YJR018W	Dn	CATGATCTGAAAGACCGCCA
YHR159W	Dn	TACTACCGGGCATGGATGAT	YJR049C	Dn	GATAGATCCATAGCTGCTTC
YHR166C	Dn	TTGCACTCGCCAGGGTCTAT	YJR051W	Dn	CAGATGGTTGAACTTGCTTC
YHR168W	Dn	TCTACTGGAGGATCTGGTAT	YJR057W	Dn	ATTTGCATAGCACCTGCGGG
YHR172W	Dn	TGATTGCCGCTCACGGAACT	YJR059W	Dn	ACTCTCAGATCAGTGTCGGG
YHR179W	Dn	TATCGGAGGGCATCCTGACT	YJR073C	Dn	GCGGAGACTTGTGTCAATAT
YHR182W	Dn	TCTCGACGTGTACTCATACT	YJR093C	Dn	GATGTGTCAGACCGACCACT
YHR197W	Dn	TCTGGCACTGGCGTTAAGCT	YJR107W	Dn	GTGGAGCTTATAGATAGCCT
YIL006W	Dn	GAGACTGCGTCATGCCTTCT	YJR108W	Dn	GAGTGTATCTCACCTATCCT
YIL008W	Dn	AATAGCTCCTCAGTGCTTCT	YJR112W	Dn	GATTCATGTCCACCTGGGCT
YIL016W	Dn	GCTACAAGTGCTGACAATGA	YKL002W	Dn	TTAGAAGAGACCGGCCCAAC
YIL018W	Dn	TGTCCACATACACGCAATGA	YKL003C	Dn	GTGATAGAATCCCATCCAAC
YIL021W	Dn	AGATGCCGCCAAGCTGTCTA	YKL004W	Dn	GACGATCAATGTCTTCCAAC
YIL023C	Dn	CCGCCTGTGCAATAATTCTA	YKL007W	Dn	GCAGAGACCATGTTAGCAAC
YIL033C	Dn	GGATATTAGCCATCTACGTG	YKL009W	Dn	GTAATGCAATTCCTCGCAAC
YIL037C	Dn	CGTGACCAGTAATATCAGAG	YKL015W	Dn	GTGCCCTAATACAGATCAAC
YIL067C	Dn	CTATGGCTACAAGGGAATGA	YKL033W	Dn	TGCCACGAAGCCATATTAAC
YIL126W	Dn	TATGTAGCTCTGCCACCATT	YKL042W	Dn	TGCCACTACAGAATGGACAC

95

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YKL044W	Dn	CGACTTGGCATCAATTACAC	YLR074C	Dn	TCTCCTAATAGAGCAGAACC
YKL065C	Dn	AACGACGTGGCAATATGCAC	YLR081W	Dn	TGAGATGTCCGAACTGACCC
YKL066W	Dn	AATAGTCGGGCAACCTGCAC	YLR084C	Dn	TATACGACGGAAGAGCAGCC
YKL073W	Dn	ATTGACCGAGAACTCATCAC	YLR086W	Dn	TACCTAACACGATATGAGCC
YKL077W	Dn	GTACCCAAATAACGGTTCAC	YLR088W	Dn	TATGGATCTACCGAAGCGCC
YKL083W	Dn	TTGGCAACAGGCTGCAAGAC	YLR115W	Dn	TTACAGATGGAAGTAGACGC
YKL096W	Dn	GGTGTTAGAACCTGCATTAC	YLR117C	Dn	TGTGAACCTACATGGACCGC
YKL099C	Dn	GGTATTCAATCCCAGCTTAC	YLR118C	Dn	TATTAGAGACCATATCCCGC
YKL102C	Dn	TAATCCACACCAGAGGTTAC	YLR131C	Dn	AAGGTCGGTCAAGTCTATCC
YKL114C	Dn	GTGCAGACATCTACAGAACC	YLR132C	Dn	AAGTGGATTCAAGGCCCTCC
YKL127W	Dn	GGCGCTTTATCAATCTAACC	YLR133W	Dn	AAGTGAGTTCAACGTCCTCC
YKL133C	Dn	TGGAGAGATACCGACACACC	YLR134W	Dn	AATATGGTCCGAACGGCTCC
YKL135C	Dn	GCTTTAACATAAGGCACACC	YLR168C	Dn	ACGTCAAACAAAGTCGTTGC
YKL148C	Dn	TCACAGACGTTGGCAAGACC	YLR195C	Dn	ATTTAGACGGACCGCGTGTC
YKL159C	Dn	CCCGATGAATTAACAGGACC	YLR210W	Dn	ACATGGGAACTGGTACTAAG
YKL168C	Dn	TGAAGGACCATCTCAATACC	YLR218C	Dn	AGACCTCTCTAACCTTTGAG
YKL171W	Dn	GGGACATAGACATTCATACC	YLR220W	Dn	AAGTAGACCTCCTCAACTAG
YKL172W	Dn	GCTCCACAGATGAAGATACC	YLR227C	Dn	AGTTCATACCCACTTCGTAG
YKL178C	Dn	ACGAGGAACGACGATCTACC	YLR228C	Dn	GCGGTAATACTACATAACGC
YKL180W	Dn	GCACATCCTTTAGAAGTACC	YLR230W	Dn	TAGAGAGAGTCCACACACGC
YKL187C	Dn	GGCCTCTAATCGAATTTACC	YLR231C	Dn	CGGAGAGACAGATTACACGC
YKL191W	Dn	ATAAAGTCGTCTCAACTGCC	YLR234W	Dn	GGAGTTTAGACACCTCACGC
YKL192C	Dn	GAAGCCTAGCGAATACTGCC	YLR264W	Dn	TAAATAGCCGTGCAACGCGC
YKL200C	Dn	TCGAGGATAAGAAGATTGCC	YLR269C	Dn	ATAGGTAGCACATCTCGCGC
YKL205W	Dn	AGATAAGATGAAGCCTTGCC	YLR270W	Dn	GTTTATACGAAAGAAGGCGC
YKL210W	Dn	ATGACCGACAGACGTTTGCC	YLR291C	Dn	AAGACAGTCCTAACATTCGC
YKL221W	Dn	GCCAGCATAGGGAATAATCC	YLR303W	Dn	TATGAAAGTCGCGTCAAGGC
YKR022C	Dn	CCCTAGACATGAAGTTATCC	YLR306W	Dn	TCCATAAGTAGCCGGAAGGC
YKR035C	Dn	GGAATACCTGTCAAGTCTCC	YLR309C	Dn	AGTATAGACCCACCTAAGGC
YKR037C	Dn	AGCCACAAAGAAGTGTCTCC	YLR310C	Dn	GCGATAAGTCTAAGTAAGGC
YKR044W	Dn	GGCATACTTTCAATACGTCC	YLR314C	Dn	ATCTACATAAGCATCCAGGC
YKR050W	Dn	CCACAGTAACTATATGGTCC	YLR321C	Dn	ATACATCCCACAGGGTAGGC
YKR054C	Dn	GAGTACCCATAACTTGGTCC	YLR327C	Dn	TAAGGATAACGACTCAGTGC
YKR057W	Dn	GATAGATCCTAACACTGTCC	YLR328W	Dn	CCTAACCTGTAATTCAGTGC
YKR059W	Dn	TGAGTACCACGACGTTGTCC	YLR329W	Dn	TAGGCAGCCACATCTAGTGC
YKR062W	Dn	GAGAAACTAACCTCGATTCC	YLR333C	Dn	CCTAATAGTTCGGAAGGTGC
YLL001W	Dn	TCATCTGTGGAACTGGCTGC	YLR336C	Dn	CATTAACCCATAAGTGGTGC
YLL018C	Dn	AGATGCGTGAGCACTACTTG	YLR362W	Dn	CATACAGCAGTGGAAGTTGC
YLL020C	Dn	ATGTGAGCTGCACTCCCTTG	YLR372W	Dn	TACAGCATACGCAGCTTTGC
YLL041C	Dn	CATTAGTGCCAAGTCGAGTA	YLR379W	Dn	GGCTAGACATGGAGACAATC
YLL047W	Dn	AAGCTGCTAGAACGCTATAC	YLR393W	Dn	ATTGCACAGCAGCCGGAATC
YLR001C	Dn	CGATATACGTGCAGACTATG	YLR401C	Dn	CGCCTGACAGTAGTATAATC
YLR010C	Dn	GTCCGCATTGCATGATGAGT	YLR402W	Dn	GGAGCGACAGCACTCTAATC
YLR028C	Dn	CACACTGTCCAATCGTAGTA	YLR407W	Dn	ACTAACGCATGGGTAACATC
YLR031W	Dn	TCTAAGTAGCAAGCACCCTA	YLR416C	Dn	GAGTCCTTACACAGTACATC
YLR034C	Dn	TCCCAGGACAGCAAGGAGTA	YLR420W	Dn	AATAGGTGTCAAGACCCGCC
YLR035C	Dn	TTGGCGAGAACATGAACGTA	YML031W	Dn	AGAGGCCCTACACTCGTATG
YLR039C	Dn	TGGCACACAGTACGCTAAAC	YML092C	Dn	CCATGCGTTTAAGTGTCTGG
YLR043C	Dn	TTGCGAAGACCAGCGAACAC	YML117W-A	Dn	AATCAGACCTCCAGTGGTGG
YLR046C	Dn	TGTTCCAGAATCGAAGACAC	YML128C	Dn	GTGACAGCTCTCAGCATTGG
YLR054C	Dn	TGGACACCATAAGACTTCAC	YMR033W	Dn	ACGCTAGAGCTTTGTAACTG

96

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YMR049C	Dn	TATGTGTTTCCCACTGCCGG	YNL137C	Dn	AGAATTACGCGCAAGTGGTA
YMR059W	Dn	TAGAGCCCTATCCATACGGG	YNL141W	Dn	AATACCGCGCAAGCGTCGTA
YMR061W	Dn	TTACGATTCCCAGGTACGGG	YNL143C	Dn	AACAGTCTCCAAGAGGCGTA
YMR063W	Dn	TATGCCGACTCATTACCGGG	YNL144C	Dn	AGAGCCTTCCAATATCCGTA
YMR065W	Dn	AACCTGGATTACAGTCCGGG	YNL146W	Dn	AAGAGCTGGCAACACCCGTA
YMR076C	Dn	GCGCATTGATTTAACATGGG	YNL149C	Dn	GCAGCCATGATATACACTTC
YMR093W	Dn	ACTCCATATCCATTTCTGGG	YNL155W	Dn	AGATAGAGTCCACGACCATC
YMR094W	Dn	CGATGCACTGACTATATTGG	YNL162W	Dn	CTATATTATGAAGGGTGCGC
YMR098C	Dn	AACTTCAGGGCAGCACTTGG	YNL166C	Dn	GCAGATCACACAGATGGTTA
YMR106C	Dn	CTAAGGGCTGTTAATCTTGG	YNL192W	Dn	ATGAGTATGCAGCCCTCCAT
YMR111C	Dn	AATGCCCAGTTCCACGTTGG	YNL208W	Dn	CCGCCTATGATACATGATTC
YMR117C	Dn	ACAGTTAAGTGCCATGTTGG	YNL223W	Dn	TCCGAGATGATAAGATAGCC
YMR119W	Dn	TTACTACGGCCCAGTGTTGG	YNL229C	Dn	GGTCCAACATCATGCCAATA
YMR126C	Dn	ATACTCGTTAGCAGGTTTGG	YNL234W	Dn	CCAGACGAGACAATCTTTCA
YMR146C	Dn	GGCAGCTATGTCGATATGCT	YNL235C	Dn	CGGCGCATAACCATATTTCA
YMR168C	Dn	CCATGTAGATCAAGTAGCAC	YNL236W	Dn	GAGAGCCCGACAACATTTCA
YMR184W	Dn	TTCTACACCATGCAGGACAG	YNL240C	Dn	GTAGTAGAAACATGCGCCCA
YMR193W	Dn	ACATCACGAGTCGATATTGG	YNL244C	Dn	CATGCAAGTACACGTTCCTA
YMR198W	Dn	AATACGGAGTTAGCAGACTG	YNL263C	Dn	TATGCTGCGCCAGGTTCGAT
YMR203W	Dn	GAGTTAGCCTCATCTATGCT	YNL264C	Dn	TATCCGACCCGCATGATGTG
YMR207C	Dn	GCTGAGGATGTACTGATCTT	YNL290W	Dn	GTGATACACTCAATCCAGAC
YMR210W	Dn	GCGTGCAGCAAACTCTACAA	YNL305C	Dn	GCTCTCTCGAAATACACGAA
YMR272C	Dn	GAATCAGCTCCAAGACGCTA	YNL307C	Dn	CCTTGCATCAAAGCCTAGAA
YMR302C	Dn	CACGCTCTGAGATTATGGAT	YNL312W	Dn	GCACTGCTCGCATTGTCGAT
YMR309C	Dn	CCGTCGAGACAATAATGGCA	YNL317W	Dn	CGCTACGATATAGAGATGTG
YMR314W	Dn	GATATGGCAACACATGCGGA	YNR001C	Dn	TAGGATTGGCCCATTCCACG
YNL011C	Dn	ATCTAGCAGTGCAGTATGTG	YNR045W	Dn	TTCATCGCACCTAAGGGACG
YNL011C	Dn	CAGACTCAGCTTTAGGAACG	YNR049C	Dn	ATCAGCCAGTCACATGGACG
YNL014W	Dn	GCCATCGCTCCATTTATGTG	YNR071C	Dn	TAGAAACACACACCCGTGGG
YNL014W	Dn	TCGGACCCAGGTTGATAACG	YNR072W	Dn	AATACCGTTTAAGCCGTGGG
YNL028W	Dn	TTTACCCAGCGGACTTAACG	YNR075W	Dn	TATTGCACCCTCCATGTGGG
YNL055C	Dn	AGGCATCGTCCATACCTGTG	YOL003C	Dn	GCATAAGCCTGAATTGCACC
YNL061W	Dn	ATATAGACAGCCCGAAGGCC	YOL040C	Dn	AATACGCGAGAATTGGCTGA
YNL068C	Dn	ACATTGATAACACGGACGCC	YOL072W	Dn	CTCTAGTATGCAGGATGTTG
YNL080C	Dn	AGAGTAAGCCACTATATCCC	YOL077C	Dn	AGGCCATGCTTCCATAGTAT
YNL081C	Dn	AGATACGTCGAACATTGCCC	YOL086C	Dn	TCACTCTCTACAGGTGTGGG
YNL083W	Dn	ATAGATGTGGAACACCGCCC	YOL098C	Dn	TTACCTATCAGATCGCATGG
YNL091W	Dn	ATAGCTGAACCAGGCGTACC	YOL101C	Dn	TGTGTCACCGCACTTCATGG
YNL094W	Dn	ACAGCCTAATAATGGTGACC	YOL102C	Dn	CGAGCACCGAGATTTCATGG
YNL100W	Dn	ATATTGGACGAATCGCCACC	YOL127W	Dn	AACCCGATGGCAGATACTGG
YNL101W	Dn	AATCCGAATCTAGTGCAACC	YOL135C	Dn	GACTGTATTTAAGTGCCTGG
YNL102W	Dn	ATGAGGAAGTCTCCGCAACC	YOL150C	Dn	AGTCTACACTTCGATCCATG
YNL106C	Dn	AGAGCTGGAGAACCTTGTAC	YOR004W	Dn	GCGTCATAGAAACCTTCACA
YNL107W	Dn	ACATGCGTAGAAGGCCGTAC	YOR046C	Dn	CCTGATAGTGCAGTCTGTAT
YNL112W	Dn	AAGGACTACGAATCTGTGAC	YOR091W	Dn	ATCCAGATTCGACGATACTG
YNL113W	Dn	AGCTACATAACCCTAGTGAC	YOR098C	Dn	GCAGATCACTCAAGCGTAAC
YNL115C	Dn	ACACCCGTAGAATCAGTGAC	YOR105W	Dn	GATTATTGCACAGTCCTCTC
YNL120C	Dn	ATAGCGTCAGAAGTCCCGAC	YOR142W	Dn	GCCGAAGTTATCGCAATGTA
YNL121C	Dn	AGAGGAGGCTCAACCTTCAC	YOR150W	Dn	GCATAGTTCGCAATCCCTTA
YNL125C	Dn	ACATACCGAGAAGTGGGCAC	YOR151C	Dn	GCTGTTCTCCAAGATCCTTA
YNL130C	Dn	AATCTTTGCCAAGCGCGTTA	YOR154W	Dn	GCGCTACTAACACAGTCTTA

97

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YOR155C	Dn	GCACGTAGAACACTCAGTTA	YPR036W	Dn	TAGTATCGACCACCGGGTTC
YOR168W	Dn	GGCTCATCCTCAAGATTAAC	YPR043W	Dn	GGACCTACACTACATTGTTC
YOR178C	Dn	GGCATACGAGAATAGCCCAC	YPR045C	Dn	GGCGACATCCTTTCAATTTC
YOR182C	Dn	GATATTCCAGAAGTCGCCAC	YPR047W	Dn	GGACCCTCGATCTTAATTTC
YOR186W	Dn	GCGCGACACTAATTTATCAC	YPR066W	Dn	TAGACGAGCCGACATCTTTC
YOR187W	Dn	GGATAGCACGAAGACCTCAC	YPR077C	Dn	CTCTAGCTTCACAGACAAAG
YOR192C	Dn	GCGGTGATAACATCTTAGAC	YPR084W	Dn	CCGCCTATTTAGACAGAAAG
YOR201C	Dn	GATGTGTACCCACCAGGATC	YPR085C	Dn	TCCGCTAAGGATTGAGAAAG
YOR202W	Dn	GACACGGATACACAGGGATC	YPR086W	Dn	GGAGCTTTAGCATCCTTGTG
YOR203W	Dn	GCCACGTATCTAATTGGATC	YPR092W	Dn	GCTCATCCATGTCACTAAAG
YOR204W	Dn	GCCTCACGTTAATTCTGATC	YPR105C	Dn	GGGTTGAACATCCTAGTAAG
YOR207C	Dn	GCGCGGATATTAGATAACTC	YPR110C	Dn	TGTTCACCCATCATCGTAAG
YOR218C	Dn	GCTTTAAGGACAGTGATCTC	YPR112C	Dn	ATTGTGAACATCCCGGTAAG
YOR219C	Dn	GATAGATTACCAGACCTCTC	YPR113W	Dn	ATACATGCGACAGCGGTAAG
YOR221C	Dn	GGCACTTAGACATACTTCTC	YPR114W	Dn	TCAGGACCATTAAGGGTAAG
YOR224C	Dn	GCCTGTATCATAACGTAGTC	YPR115W	Dn	CCGATGCGTATAATGGTAAG
YOR233W	Dn	GCACGCACGTAACTTGATTC	YPR119W	Dn	CAGTCTGAACGCAGTGTAAG
YOR238W	Dn	GGACCAGCGTCATAATCTTC	YPR121W	Dn	AGTTCGACCCACGCATTAAG
YOR249C	Dn	GCGCAGTAATGCTTTAGAAG	YPR131C	Dn	AACCATACGGGTGTGAACAG
YOR258W	Dn	GCTATACACTACTAGACCAG	YPR157W	Dn	CCAGGAGTATAATGTGACAG
YOR260W	Dn	GATCAGATACCACTTTCCAG	YPR159W	Dn	GCACGAGTATTAACCTACAG
YOR278W	Dn	GCTACAGTACGATCACCTAG	YPR161C	Dn	AACTTTAACGCACCGTACAG
YOR281C	Dn	GTAGCTGTAAGCACTGTTAG	YPR165W	Dn	ACAGGCCACTTAACTTACAG
YOR353C	Dn	CATCTGTAGGAAGTAGTAGC	YPR170C	Dn	TAGTTATACAGCCCGACCAG
YOR367W	Dn	TCGCATGAGGCATTAGTATG	YPR173C	Dn	GAGTAATGAGCATGTACCAG
YOR368W	Dn	TATCATGCCGCAGGCGTATG	YPR175W	Dn	AGAGCGATTACACGACCCAG
YPL024W	Dn	TAGACTCGCTACATCCTGGG	YPR180W	Dn	GGATGCTACCTAATCGCCAG
YPL065W	Dn	TATGTCCACCGATAGCCAGG	YPR181C	Dn	TGAACTAGCCGTAAGGCCAG
YPL146C	Dn	CGATTCATTCGCATTGACTG	YPR183W	Dn	AAACCTTTCTGAAGTGCCAG
YPL157W	Dn	ACTGTTCCAAGAGCTGCTTA	YPR193C	Dn	ATTGAAGTACGGCTAAGCAG
YPL160W	Dn	CACTAAGGCCAATCACTTGA	YPR195C	Dn	AACCTGACCTAATGCAGCAG
YPL177C	Dn	CAGCCGAGAGTGTGATACAT	YAL025C	Up	TGCCGCATCAAAGAGGCCAA
YPL180W	Dn	AATCTCGCATTCCTGGACAT	YAL034W-A	Up	GGACCTCTGCTCATTATGCT
YPL195W	Dn	TATGAGCCAGCAGACCTGTC	YAL044C	Up	ATCAGGTCACGCAGTATTGG
YPL197C	Dn	GTATATCTCAGACCACTGTC	YAL047C	Up	GTCCGACGTTAGATCACCTG
YPL208W	Dn	CTATGAAGGGAATGCACAGC	YAR008W	Up	CATGAGAGTGAAGCAGTATC
YPL209C	Dn	TCTAGGACAGCATACACAGC	YAR019C	Up	GCGCTTATCACATTTGACAG
YPL218W	Dn	CCCGATATGCAAGGATACTA	YAR042W	Up	ATTCATGTGCCAGTGCCGTG
YPL232W	Dn	TAGATCGTGATGACGTTGCT	YAR043C	Up	ATTCTAGCGGCAGATCCGTG
YPL234C	Dn	TATGTTGCCGCATCTCCGAT	YBL021C	Up	ATAGTGACGGAACTGTTAGC
YPL255W	Dn	GAGCTAACGCTAGATATGTC	YBL023C	Up	AAAGGAATCTGTCTCAACGC
YPL260W	Dn	ATACGTCCTGAATAGCACTC	YBL046W	Up	ACAGCAAAGGAAGTTGTCGC
YPL262W	Dn	CTCAATATGTGGGAGAATGC	YBL049W	Up	AACGTATCGGAAGCATAGGC
YPL268W	Dn	CGGCAGTATGAATAGTAAGC	YBL058W	Up	AAGATACCGTAACATGCGGC
YPL274W	Dn	CCAGCTAGGCAATCATACTA	YBR020W	Up	TATCCCTACGGCATTGCGTG
YPR005C	Dn	TTATATGGCCGCACCCGATG	YBR101C	Up	ACTATACGAGGTGGCTGAAT
YPR016C	Dn	GTAACACGGTTCTGAAGTTC	YBR102C	Up	ACGCGACGACTTCTCTGAAT
YPR027C	Dn	AGAGTTACCTAATCCCGTTC	YBR108W	Up	ACTTCTGCCACGGGTGACAT
YPR033C	Dn	GTATATCAGCCCATTCGTTC	YBR121C	Up	AGGATCGCGTAGACGTTCAT
YPR034W	Dn	GGTTATACCCTACCAGGTTC	YBR135W	Up	AAGCTATTCGCATTCGGGAT
YPR035W	Dn	CACAGTGGATAACTAGGTTC	YBR139W	Up	ACTGGGTACTCATGTTGGAT

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YBR142W	Up	ATTATCGGCCCACGGCTGAT	YDL126C	Up	CCGACCCTATTAGCTGATAT
YBR170C	Up	ACTCCCGATGTATTGACACT	YDL129W	Up	CTATGTGCGGTAAGACGTAT
YBR193C	Up	AGACTGTCGTCAGATCCGGT	YDL134C-A	Up	CCTAGCCTGAGAGGGATTAT
YBR213W	Up	ATATCGCCCTGAGGACCTGT	YDL135C	Up	CGCGAGCTAGGCGTACTTAT
YBR221C	Up	ATGTACTCCCTCTCAGGTGT	YDL140C	Up	CAGGGATATTGACTACGACT
YBR223C	Up	AGACTCTCCTTAGACTGTGT	YDL141W	Up	CATGTACGAGTAGTAGGACT
YBR238C	Up	ATGAGTCTCTTCCACCGATT	YDL142C	Up	CCCGTTCATTCATAGTGACT
YBR243C	Up	ACGTAGGGATGATCGCTATT	YDL146W	Up	CTCTAGTAGCGGAGATACCT
YBR248C	Up	AGGCACTTGCTCCAGGACTT	YDL163W	Up	CACTGTCTACGATGGGTTCT
YBR272C	Up	GAGATTATCGCATACGCCTG	YDL187C	Up	CCATTACTGTAGATGACGGT
YBR277C	Up	CAATGACGAGTTGAGGCAAT	YDL195W	Up	GCCTTAGCCAAATAGGGCAA
YBR282W	Up	TGCTACGAGCTATACTCACT	YDL196W	Up	GGGACCGCCAAAGCTATCAA
YBR295W	Up	CCAGCTACTAAAGGATGTCA	YDL206W	Up	GCCCTCACGAAATAGTTGAA
YBR300C	Up	GAGACCAGACCAACCTGTGA	YDL207W	Up	GCCCGAACCACAATGTTGAA
YCL022C	Up	CCTAGCGGTAAAGTGATAGA	YDL208W	Up	GGGTGCCACGCCAAACATAA
YCL024W	Up	CCCAGGAGTAAATCGCTAGA	YDL216C	Up	GCCCTGATAACAAGGTGTAA
YCL026C	Up	CCCTCTGCTAAAGTAGTAGA	YDL220C	Up	GGCGTTCCTAAACATCAACA
YCL031C	Up	CCTCATAGTAAAGTCACCGA	YDL221W	Up	GGCCGTCATAAACGCGAACA
YCL034W	Up	CAGAATGGTAAATCCCGCGA	YDL223C	Up	GCTCCCTTTGAAGAAACACA
YCL036W	Up	CACCATTTGAAACGGATCGA	YDL233W	Up	GGGATCACAACCACGTTACA
YCL043C	Up	CATTCGGGTAAATCTGAGGA	YDL234C	Up	GCCGTACACACAAGGTTACA
YCL059C	Up	CCGACTGGACAAGTTAATGA	YDL236W	Up	GGCGTCAAGACAACGTACCA
YCL060C	Up	CGCCCGAACGGAAGATATGA	YDL238C	Up	GCCACTTAGAAATTAGCCCA
YCR014C	Up	CCTAAGTCTCAAGAAGGCTA	YDL239C	Up	GGCCAGAAGAAATCGTCCCA
YCR026C	Up	AGCAGAGTCTAACAGATCCC	YDR002W	Up	GGCTACCGCAAATAATTCCA
YCR052W	Up	AAGCGTCCATGTGCCTCAAT	YDR027C	Up	GGCCCTAATCTAAAGCCAGA
YCR057C	Up	CTCTCATCGAGGGCAGGAAT	YDR035W	Up	GGAGTACCAACATATCCCGA
YCR061W	Up	GGTGCTCAGTCATCTTCACT	YDR047W	Up	GATAATTCCCAACCATCGGA
YCR073C	Up	GCCTATGTCACAATCTGTCA	YDR050C	Up	CCTCTGAGTAAATACGGAGA
YCR073W-A	Up	GATAACCAACCAGTCTGTCA	YDR050C	Up	GTACATAACTTCAAGCGACC
YCR077C	Up	CAGCGAGCAACACTCTGTGA	YDR052C	Up	GCTATCTACGAATTAGGACC
YCR095C	Up	CATGGTAGATAACTGGCATC	YDR054C	Up	GAAGCGTGACGAATCTTACC
YCR102W-A	Up	AGGGACCAGTCACAGTCATC	YDR060W	Up	GGATCACCATAATAGTACCC
YDL021W	Up	CCACGTTGGTCAATATGGGC	YDR069C	Up	GACTCTATTCAAGGTTCTCC
YDL026W	Up	CCAGGACACTAAGGTAATGC	YDR075W	Up	GAAGCAATAACAGCCGTTCC
YDL030W	Up	CCGGACATACTTAATTCTGC	YDR082W	Up	GATAGACCTTAATTCACCGC
YDL034W	Up	CCTGAGTAATAAGTCCTTGC	YDR084C	Up	GGCCAACAAATAACTTGCGC
YDL036C	Up	CACAGGAGGTAACACTTTGC	YDR086C	Up	GCGGTTAATAGACATTTCGC
YDL042C	Up	CCCTGTTATGAACCTTGATC	YDR090C	Up	GAGACTACTGAACCTTCGGC
YDL065C	Up	CCCTAGAGAGATTTCTGAAG	YDR091C	Up	GAGATTTACTAACCCTCTGC
YDL066W	Up	CCGCTAAGACTGTATTGAAG	YDR092W	Up	GACTACCTAATACGACGTGC
YDL069C	Up	CGCGTAAGAGTATAGTACAG	YDR102C	Up	GCCACCGCTTAATTTAGATC
YDL075W	Up	CCATTGGACTAATACGTCAG	YDR103W	Up	GGACACTCTTAATTCCGATC
YDL077C	Up	CCGTGTCTATAAGTGTTCAG	YDR105C	Up	GTACCTCGTAACATTCGATC
YDL097C	Up	CGACCTCTTACAGTGATTTG	YDR113C	Up	GGTTTGACCACCTATATCTC
YDL098C	Up	CTCACTTGGAGAGGTATTTG	YDR117C	Up	GAAGCACCTTTCACGAAGTC
YDL102W	Up	CTTACGTCAGGCGTGGCAAT	YDR147W	Up	GCTCACTTGTTACAGGTACT
YDL103C	Up	CGGACGAGCTTCCATTCAAT	YDR158W	Up	GACGTTTAGGCACTACTGCT
YDL105W	Up	CCGCGACGATTGATTAGAAT	YDR168W	Up	GAGACGCCGGTCATTCTTCT
YDL108W	Up	CAGAGGGCACTGTTCTTAAT	YDR170C	Up	GTAACCGAGTGTCTATCAGT
YDL120W	Up	CCGAGCTACGGATATTTGAT	YDR173C	Up	GCGATTCTGGTACATTACGT

99

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YDR176W	Up	GATTATACGCTATCCGAGGT	YDR518W	Up	AGCAGTGGCCGATTCCCTTT
YDR179C	Up	GGAGCTTCCCTCATCTTGGT	YEL017C-A	Up	CGACTACAGGCATATTCATC
YDR179W-A	Up	GCAGTTCATAGACCTTTGGT	YEL026W	Up	TCCCGCATTCCAGATGATGG
YDR183W	Up	GCCGCTACCTTGACTGAATT	YEL032W	Up	TCATCGGACTCACGGTGCAT
YDR187C	Up	GCGAGGCGTATAGTTTCATT	YEL035C	Up	AAGATGCTTGGACACTAGCT
YDR188W	Up	GCGCCTTAGTTTCACAGATT	YEL041W	Up	CCAGCATTGAAATTCTGCCA
YDR189W	Up	GGTCGGACTCTATACTGATT	YEL045C	Up	TCCTATGAGACAATGGGAGA
YDR193W	Up	GTGATCCGGCTGCCTAACTT	YEL046C	Up	GCGCCATCGAACCAATGAGA
YDR200C	Up	GAGCGTAGCCTTTCATCCTT	YEL055C	Up	CGCTGACAGTAACTTTCATC
YDR211W	Up	GAAATTCGCGTTCATGCGTT	YER006W	Up	GAGACTGCTAAACTCTGAGA
YDR217C	Up	GAGCCCTGCTTGGTCACTTT	YER007W	Up	CACATCTAGCCAAGGTCATA
YDR225W	Up	TGCTCTACCAAAGCCGTAAA	YER008C	Up	GCACCAGAGCAACTGTCATA
YDR226W	Up	TGGCCCTCAAACCATGTAAA	YER031C	Up	TTAGCGCACAGCCTGACAAG
YDR227W	Up	TGCGCCCTAAACAGCTTAAA	YER036C	Up	CACGCCCTTACATGATATGG
YDR237W	Up	TCTCAGCCGAAAGAGGGTAA	YER039C	Up	ATACAGCCTGGCACAGTCTG
YDR266C	Up	TTAATACGGATGCCCAGAGG	YER047C	Up	ATATTGTCTCACGCGGCGCT
YDR268W	Up	TATTCACGTAGACGGATAGG	YER051W	Up	TGGGCACTCACAATAATCCA
YDR269C	Up	TTTATGCGCCCAGGACTAGG	YER073W	Up	ACATGAGCCATAAGTGTCTC
YDR270W	Up	TTCGCACTGACCGTACTAGG	YER091C-A	Up	GAGGTATTGGAATCCTGATC
YDR279W	Up	TATACGACCGCAGGATTCGG	YER122C	Up	CATTGCAGTGAAGTGAGATC
YDR280W	Up	TAAACACACTCCGCAGAGGG	YER125W	Up	GCTCAGCAGTAATCATTCTC
YDR283C	Up	TATCGTACTCCATGACCGGG	YER126C	Up	CATAGGCTTGAAGGATTCTC
YDR284C	Up	TTACTTGTGACCATGCCGGG	YER139C	Up	ATATGCGGCTACATCCCTGG
YDR300C	Up	TATCAGCTCCCAGGGTTTGG	YER147C	Up	CGCCCTGGATCATTTATCAT
YDR301W	Up	TTCCCTAGTCCAGACGGATG	YER152C	Up	GCTATTAGTGCATCATCGCT
YDR303C	Up	TTCACGGTACGCACGGTATG	YER162C	Up	CGCCCAATGCAATTTGTAGA
YDR308C	Up	TAGGACGTTAGAGACTTCTG	YFL010W-A	Up	CGACGTAATCCATCAGGCAG
YDR315C	Up	TTTCACTTCCGGCATGGGTG	YFL012W	Up	CTATATGGCTAATGAGGCAG
YDR318W	Up	TTTACCCGGCGCAGTGATTG	YFL018W-A	Up	TATAGACGCCGACCATGCAG
YDR325W	Up	TTCTGTACGCCACCGTGTTG	YFL020C	Up	AGCTCACACCTAATATGCAG
YDR328C	Up	TCGGACGACGGATTTGACAT	YFL023W	Up	ATATAGCTCCCACATTGCAG
YDR335W	Up	TTAGGCCCGAGACTCCTGAT	YFL035C	Up	CCCGTAGTGATAATACTCAG
YDR337W	Up	TTTCGGATACGAGTGGCTAT	YFR005C	Up	GGACCCAGTTACCATCAGAG
YDR355C	Up	ATATCACGGCGGTAATCCCG	YFR019W	Up	CATTGCATGTAAGGCTAGAG
YDR376W	Up	AACACGCTTTAACTACTGCG	YFR048W	Up	CCGTCATAGACATGCTTATC
YDR392W	Up	ATTACCTAGTCGGAGACAGG	YFR050C	Up	CCGCAGCTCTAATTGTTATC
YDR398W	Up	AACCAGCACTTGTCTAACGG	YFR055W	Up	AGAGGCCATCTCGTGAACTC
YDR401W	Up	AAATTATTCTCACGCACCGG	YFR057W	Up	CCACCAGGGTAATTTAACTC
YDR408C	Up	ATTCTTCTCACACGGGTGGG	YGL001C	Up	CCTTTGGGAGAAGTGAATAC
YDR410C	Up	ATGCACCGTCCCTAATTGGG	YGL002W	Up	CCCAGGGTAGAATAGGATAC
YDR418W	Up	ACTACGGTTACAGGAGTATG	YGL003C	Up	CCAGGATTAGAACCTACTAC
YDR433W	Up	ACGCTAGACTGTTGGGAAAT	YGL006W	Up	CCTGATTTAGAAGAGGGTAC
YDR448W	Up	AACGCCTTCTCAGTATGGGT	YGL015C	Up	CTTCAAAGATAAGTGGACCC
YDR464W	Up	AGGGCCTCCTTCTCATAATT	YGL016W	Up	CGCTTCAAGGTAATTTAGCC
YDR472W	Up	AGTGTACCTCTACTCCGATT	YGL018C	Up	CCGGGATATGAATTAAGTCC
YDR473C	Up	ACCTCGCGTATAGGGTGATT	YGL028C	Up	CATGTTAGGGAACTTTAGGC
YDR474C	Up	ACGGAGGCGGTCATTATATT	YGL032C	Up	CTTAAAGACGCGGGAACTGC
YDR483W	Up	ACGGTTCTATCAGTGACCTT	YGL041C	Up	CCAACAAACTAAGGTGTGTC
YDR485C	Up	ATTACGCGGCGAGTAGCCTT	YGL044C	Up	CCAGTTCGACTAAGGGATTC
YDR506C	Up	AACACCTACTGATCGTGGTT	YGL047W	Up	CCCGTAGACGTAATGATTTC
YDR507C	Up	ACCCGTCGTGTCACTATGTT	YGL048C	Up	CCGTACCTGGAATATCTTTC

100

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YGL073W	Up	CGGGTGTAATACTATATCCG	YGR081C	Up	CTGGACTAGGTACTTAGGTT
YGL092W	Up	CTACGTGGTTAAGTACGATG	YGR083C	Up	CCCTGGGAGTTAGTAGATTT
YGL098W	Up	GCTGCCCGGAAATAACCGAA	YGR085C	Up	CCCGTTAGCGTATGGGATTT
YGL099W	Up	GCCTTTCTAACAAGACCGAA	YGR086C	Up	CACGGCGGTTAGAGTTATTT
YGL102C	Up	GCCTACCGAACAACTTTGAA	YGR091W	Up	GTGGCCGTAACCAAACCGAA
YGL103W	Up	GCCGGTTACACAACGTATAA	YGR094W	Up	GCCTCCGGTAAAGGAAACCA
YGL104C	Up	GCCGGGACGAAATTCACTAA	YGR096W	Up	GCCCTTAATCAACGGGTTTA
YGL105W	Up	GCTTCCTAGAAAGGGACTAA	YGR099W	Up	GATTAACCGTAAGGGACTTC
YGL109W	Up	GGGCACCCGAAATACTCTAA	YGR108W	Up	GGCGGAGTTAGTTCCAAATT
YGL111W	Up	GGCCCAGTAACAACTCTTAA	YGR110W	Up	GGAGGCTTTACTTAGACGTT
YGL116W	Up	GGTGCCCTTAAATACAACCA	YGR112W	Up	GGGTCCTACCGCCTATATTT
YGL123W	Up	GGTGGCTACAACCCAAACGA	YGR119C	Up	TTACTTTAGCCCACCGAGGG
YGL125W	Up	GTTACAACCAAACTGGCCGA	YGR120C	Up	TTACTACCGCCAGGTTAGGG
YGL128C	Up	GGAGTAACCCAATCCACGGA	YGR135W	Up	CGGATATGGACAATCATCGA
YGL129C	Up	GCTCCCGTTAAACCTTATGA	YGR138C	Up	CCAGGTTAAACACTTGTCGA
YGL130W	Up	GTACGTTAAACATCCGGTGA	YGR154C	Up	CGCCCTATTATACAATCGGA
YGL136C	Up	GGTTTATCCCAATCGACCTA	YGR156W	Up	GATCATCCAAGACACTCGGA
YGL139W	Up	GGCCGGGCAATACGAAATTA	YGR159C	Up	TCTCTAAGAAGACATCGGGA
YGL140C	Up	GCGCGAACTTTACCCAATTA	YGR166W	Up	GGCACACGATACCAACTGGA
YGL141W	Up	GCTCTAGGTAAACACCCTTA	YGR181W	Up	GGACATCATACCAAGCCTGA
YGL142C	Up	GGCCCTACTCTCAAAGTTTA	YGR186W	Up	CCGACATATACCAAGGCTGA
YGL143C	Up	GGTCAACCAGTCCTGTAAAC	YGR204W	Up	CCTACAGCAACGCGAATTGA
YGL144C	Up	GCTCCCACAGTAAGTTAAAC	YGR213C	Up	CCACTTGAAGCCAATCTTGA
YGL145W	Up	GGAGGCCACGAACTACTAAC	YGR217W	Up	TGGCAGAGAAGAACCTTTGA
YGL160W	Up	GGGACATTCCTAATAACGCC	YGR218W	Up	CCAGTATCCCAAGAGTTTGA
YGL165C	Up	GTACACGACCTAATTGTTCC	YGR219W	Up	TGAAGGGACCTCAGACACTC
YGL171W	Up	GTACAAAGATAACGGGCGTC	YGR226C	Up	AAGGTCTCGGAATATGACTC
YGL180W	Up	GAATAGGCGGTAACTCTTAG	YGR228W	Up	GGCACCTCTGAATTATACTC
YGL195W	Up	TAGGTGAGTCCACGTACCCT	YGR233C	Up	AAAGGCAGTGGAATGTACTC
YGL200C	Up	TTTCTAGCCCGAGTGGGTCT	YGR237C	Up	TCAGGAGTGGAATGCACCTC
YGL202W	Up	TTAACCGACTGTGCCTAAGT	YGR240C	Up	TAAGCCTCGATAAGGACCTC
YGL203C	Up	TCTTAACCGCGATTCGTAGT	YGR264C	Up	CCAGCAGGTTATCTAATACG
YGL206C	Up	TAACTCGGCCTAGTGGACGT	YGR274C	Up	TGGATCTTCAGACACCTACG
YGL224C	Up	TGAGTACCCGTCCGGTCATT	YHL006C	Up	GGACCCATGAAATCCTAGCA
YGL226W	Up	TCCTTTCGGCGATAGGGATT	YHL013C	Up	CAGCTCTGGCGACTATGGTT
YGL227W	Up	TACTACCCTGGAGTGGGATT	YHL016C	Up	TCTGAGATGGCACGTCTGCT
YGL229C	Up	TTGGGTTCACGGCTCCACTT	YHL029C	Up	AAAGCGCAGTGCTCATAATG
YGL231C	Up	TGGTAGACGGTCCTCACCTT	YHL037C	Up	CAGTCGAGTACATGCACCAG
YGL233W	Up	TACCTTTAGCGCGGGAGCTT	YHR005C-A	Up	CAGCGATACGCATGTGTACG
YGL237C	Up	TATTAGTGCGGACCCGGCTT	YHR018C	Up	GCACGTACCACAGATGTATC
YGL240W	Up	TTCTACCGTCGGAGGCAGTT	YHR019C	Up	CGGCCTATCAGAGAATTATC
YGL241W	Up	TTACCTTGCGGGCCAGAGTT	YHR020W	Up	CATGCAGGTGAAGGATTATC
YGL250W	Up	TGATTCCCTCGACCGTGGTT	YHR021W-A	Up	GCGCTTTAACGATTTGTACG
YGL255W	Up	AACCTCTGTAAACTTCGGGA	YHR030C	Up	ATATAGTCGCCCATGCAGCG
YGR037C	Up	CACTTTAAGGAAGGGTAAGC	YHR036W	Up	TCAGCTCAGAGGGCTTAATG
YGR041W	Up	CCCGTAGACCTAAGTTATTC	YHR040W	Up	TATCAGTGCCCATCGTCGTG
YGR054W	Up	CCCTTGTAGTAATGGGTTAG	YHR044C	Up	CCTTCTTGAGGATGTGAGAT
YGR059W	Up	CCCTAGTTCTTAATCTACGG	YHR045W	Up	ACAGCGTTATCACTTCTGCT
YGR060W	Up	CCTTTATGTTAAGTAGGCGG	YHR059W	Up	CCATCTGTGAAAGTACGTCA
YGR062C	Up	CCCTATTTGTTAAGCGTGGG	YHR069C	Up	TGCAGAGAATAACTGTGACC
YGR072W	Up	CCGGAAGGTGTTGTAACTAT	YHR072W	Up	GCACAGAGCACAGTAGTACC

101

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YHR081W	Up	TACGTCAGCACAGCCTTGAG	YJL057C	Up	ACATAGTGACCATCTCTCAG
YHR084W	Up	CACGAGTGCTAAGATTTGAG	YJL059W	Up	TATTAGCCACCATCGTATGC
YHR085W	Up	GCGATCATTAGATATGCAGG	YJL075C	Up	CGATCAGTACCAACGATGGA
YHR105W	Up	GCCCGCATGAAATGAGGACA	YJL076W	Up	CCTAGATCGCAATAGATGGA
YHR146W	Up	TCCTGTACCACGTCGATATG	YJL090C	Up	GCAGTCTACTCATTCCTGAT
YHR147C	Up	TACTCGGCACGACGGATATG	YJL093C	Up	CTCTTATGAGGGAGACTGAT
YHR152W	Up	TACTATCCGAGAGGTGTATG	YJL111W	Up	AATCTGGCGGAAGCCTATGC
YHR183W	Up	TTTATCCTGGCACGCTGGTG	YJL115W	Up	AGGCACACGGCACTATATGC
YHR197W	Up	TATAGCCGCGTCCACTGTTG	YJL116C	Up	GAGACCACATCTAATTCAGC
YHR204W	Up	TATACGCGAGTCAGGTGAAT	YJL125C	Up	GCTGACACTCAACGAATCTA
YIL004C	Up	CCCGCACAGAAATGCTTGAA	YJL131C	Up	AAGAGAGCGCAATTCTGCTA
YIL005W	Up	GCTGTCCCGCAATACAATAA	YJL140W	Up	ACTATGCTGAAAGTGTGGCA
YIL014W	Up	GCAGGAATAACAGTTCTGGA	YJL149W	Up	ACTACTACGTGCGGCATGAT
YIL015W	Up	GAGTATCGCCAACATCTCTA	YJL151C	Up	CCCACTGATCTTGACATGAT
YIL016W	Up	CACAGCCGAAGAAGTCTCTA	YJL152W	Up	ACTAGCGTCGCACTGTGGAT
YIL018W	Up	GCAGGATCAACACGAGTCTA	YJL163C	Up	ACGCACGGGTTTAAGTCCTG
YIL021W	Up	CCGATCTTATCAACGATGAC	YJL167W	Up	GATACCGCTACATTGTGCAG
YIL027C	Up	CATACGACAGGAGACACTGC	YJL172W	Up	CAGATAGGCACAGATGATGC
YIL028W	Up	CAGATAGTAGCATACACTGC	YJL181W	Up	AAGACGTGCCAACTATGCTA
YIL031W	Up	ATGACGCCAGCAGTCTTCAG	YJL184W	Up	ATTTCGGACGCCAGTTCCTG
YIL039W	Up	TACGATCCGGCATGTCGATG	YJL189W	Up	CTCCATACAACAGAGTGGCA
YIL044C	Up	TTAGAGCAGCGCGTCCATTG	YJL191W	Up	TAATTCCGCACAATCTGGCA
YIL048W	Up	TATCTCCTCGCAGCGGTGAT	YJL192C	Up	CGCATGGACAAACACTGGCA
YIL058W	Up	TTTACCCGCGCATTGGAGTG	YJL196C	Up	GCTCATGTCAAACGACCGAA
YIL061C	Up	CCAGAACCCAGAATTGTGCA	YJL200C	Up	TAGTATGCGTCACTCGCTCT
YIL062C	Up	CAATCTCGACCAATCGTGGA	YJL206C	Up	TAGTCGTCAGCACGTCTGTG
YIL064W	Up	CAGACTGCAACATACTTGGA	YJR010W	Up	CCGAAGCACAGACAGTGTTA
YIL065C	Up	GCTACACACACAGAGTTGGA	YJR012C	Up	ATGACAGGTGAAGACACGCC
YIL077C	Up	AGTATCACTACATGCACTGC	YJR013W	Up	AGATAGATGAGACCCACGCC
YIL083C	Up	CATGACCATCTAGTACAGAG	YJR015W	Up	AGCAGACTCGAATGTTGCGC
YIL092W	Up	TTTGCACCTCGGCCATAGTG	YJR017C	Up	GATTACATAACCTGCATCGC
YIL110W	Up	GCCCAGATGAAACTAGGCCA	YJR021C	Up	AATTGGCGCGAATGCCCATC
YIL113W	Up	GCGCTTAACCAACTGACAGA	YJR030C	Up	CAGCCTAGTACATTAGCTTC
YIL115C	Up	TCCCTGAATACAATGCCAGA	YJR035W	Up	ATATCATAGCCCACTGTCGG
YIL117C	Up	CCATCGCAGCAACTGTAATA	YJR036C	Up	TCATGCTTAGCCAGATTCGG
YIL118W	Up	GGCAGCTATCCAATCACATA	YJR046W	Up	TTCAGGTACGGCATGAGCAT
YIL126W	Up	CAATCTAGGCGAATGAGGCC	YJR047C	Up	GATATTGATCTCACCTGCCT
YIL143C	Up	AGATTGACTACACGCTCTTC	YJR057W	Up	CATCTCCGACAAGCTGAAGA
YIL148W	Up	CACTCTGTTTGAATGATGGG	YJR065C	Up	CCAGCAGCTCAATCTTGTTA
YIL150C	Up	TACTGCACCGCATCTATGGG	YJR068W	Up	GACTGACGTACATTGTGACG
YIL153W	Up	TACGATTGCAGAGATTGCTG	YJR107W	Up	GATCACGTCTTCTAACACTG
YIR003W	Up	ATAGCATGAACATCACGGCC	YJR108W	Up	GGTGTATCCAGCTACCACTG
YJL006C	Up	TGCATCTAGTGTCAGAGTGT	YJR123W	Up	GCTCCGATAGGATATTAGTG
YJL008C	Up	CAGCCGATGTTCCAGTCTGT	YJR125C	Up	GAGACTATTCCATACTCGTG
YJL017W	Up	TACTATGTCCCATCTGGCTG	YJR141W	Up	GCTGCGTCCATTTGAACAAT
YJL018W	Up	TACTGTCTGAGCACTGGCTG	YJR149W	Up	GCAGTCGCAGTCGCCTTAAT
YJL025W	Up	CCATGTCTGCTAATGTGTAG	YKL004W	Up	CCCTAAACAACAGGTTCGTA
YJL026W	Up	ATACTAGCAGCACGGTGTAG	YKL006C-A	Up	CCTCGGACAGAATATAGGTA
YJL028W	Up	AGCACTAGCCCATTCTGTAG	YKL009W	Up	TACAGACCCACATACTGGTA
YJL029C	Up	AGATCGACTCACAGGACTTC	YKL018W	Up	CCGAGTAGAATCAAGGTGTA
YJL039C	Up	TTTGCATCGCGCACGCTGAT	YKL019W	Up	TAGACAGAACCACGGGTGTA

102

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YKL021C	Up	AGACAGAGGCAACTTGTGTA	YLR065C	Up	TCGCAATGCAAAGCCGGGAA
YKL024C	Up	CCGTCGAACACACATTTGTA	YLR066W	Up	TGGAGCCTCAAATCCCATAA
YKL028W	Up	CGCGGAACACCATTCAATTA	YLR071C	Up	TCTTGATAGAAAGGCGGACA
YKL033W	Up	AGCCCGTCAAGAGTAACTTA	YLR075W	Up	TGGGTATGCACAACCAACCA
YKL060C	Up	CCGATATGTCGCAACCTTTA	YLR076C	Up	TTGGCGGAGACAACACACCA
YKL065C	Up	CGGCTTAACAGAACCGTTTA	YLR086W	Up	TCCTGCCGGAGAAGAAAGCA
YKL073W	Up	CGCCTCGATATGAATCAAAC	YLR088W	Up	TTGCCGGGAGACAAACAGCA
YKL086W	Up	GGCCAGCAGATATGTTAAAC	YLR092W	Up	TCCGACGCAACAATAGGGCA
YKL108W	Up	CTCCGATAAGCGAAGAGGAC	YLR105C	Up	TCCCGAATGACAAGGCACGA
YKL138C	Up	GACACACAGGTTTCCAATAC	YLR115W	Up	TCTGCGAGCCCAAGAAAGGA
YKL139W	Up	TTTGCATCGGCATCACGCTG	YLR116W	Up	TCAGGCTGTAAACTGCCGGA
YKL150W	Up	CGGGCGACATAAGCAGATAC	YLR117C	Up	TCCCAGCGAAGAATATCGGA
YKL152C	Up	CGACGAACCGAATGCGATAC	YLR124W	Up	TTGTGCCAGACACCCAAATA
YKL154W	Up	CCGACACTGTGAACATATAC	YLR125W	Up	TGACCCAAGCTAGTCCAATA
YKL159C	Up	GGCCACATAGCAGGAACTAC	YLR127C	Up	TGGGATCTTCTCACCCGTGT
YKL163W	Up	GCTCGACATTAACAGACTAC	YLR129W	Up	TCTATCCCGCTAGGGTGTGT
YKL166C	Up	TGAAGAGGCACCGAACCTAC	YLR134W	Up	TTACGGAGGCTTGGCATATT
YKL169C	Up	GCGATCTAAGTCAATGCTAC	YLR135W	Up	TCCTCGGAGCTAGGGCTATT
YKL171W	Up	GGCTTCTCAACATTTGCTAC	YLR147C	Up	TTAACCGTCTGGAGATGCTT
YKL176C	Up	TGGCAATCGTCCGCAAGTAC	YLR190W	Up	AGAGTCGCCCAACCGTTATA
YKL180W	Up	GCCCATTACTTAGAACGTAC	YLR202C	Up	ATACCGCCAGAAGGGTTCAC
YKL186C	Up	AGACATCCCGAATCTGGTAC	YLR206W	Up	AGACGACTAACACCTTTGAC
YKL197C	Up	TGGAGATAGACAGACCACCC	YLR232W	Up	GATACGAGCACAGCAGTTCC
YKL200C	Up	AGGAGTAACTTCAACGACCC	YLR234W	Up	GGATTATCACCATACGTTCC
YKL218C	Up	GTAGATATTACACACTGCCC	YLR238W	Up	CGGCGCTAATACTAATTTCC
YKR001C	Up	GGTCCATATTTAGCAATCCC	YLR244C	Up	CCGCTCTACTATAAGAAAGC
YKR006C	Up	GTACTTAGTCAATTCGTCCC	YLR262C	Up	CTTGACAGAGGACATGAAGC
YKR025W	Up	AGGACATGGACACACTAGCC	YLR269C	Up	GTCCACCATATAACGTAAGC
YKR031C	Up	TATAAAGATTGGGCAACGCC	YLR275W	Up	TGAATACATGGGAAGCCAGC
YKR034W	Up	TATGATGAGACACCGACGCC	YLR310C	Up	CTAACACTGGTTCAAGTAGC
YKR095W	Up	AACCAGAGTTTCTAACCGCG	YLR321C	Up	TTTGGAAGACTCCCGAACGC
YLL018C	Up	ATGGCTCATCGCCTCAGAGT	YLR322W	Up	GGTATAATAGCATGGAACGC
YLL021W	Up	GCCCTGCTGGAAATCAAACA	YLR329W	Up	ACTAAGATGACATATCCGGC
YLL023C	Up	GCTCGGCATAAATCTCAACA	YLR330W	Up	CATACATCATAAGTTCCGGC
YLL028W	Up	CGCGCTAAGACAATTCATCA	YLR332W	Up	TCATTCCTAACAGTGGCGGC
YLL029W	Up	CCAGATGCTACAATCCATGA	YLR333C	Up	CATCTGAACCGATATGCGGC
YLL036C	Up	ATAGCTCTCCAACGGCAGTA	YLR334C	Up	CCATGCTAGATAAGTCGATC
YLL037W	Up	CAGGCCATAATGCTCAATAC	YLR363C	Up	AACCATTAAGCAGACGTGGC
YLL038C	Up	GGATGCACACACTTCAATAC	YLR364W	Up	TCCATGTATATGAAGGTGGC
YLL040C	Up	GTGGTACACAGCACACATAC	YLR365W	Up	TTATATCAGGACCCGGTGGC
YLR006C	Up	GCTGCGGCTAAACCAGAACA	YLR366W	Up	CGGATATTTCTCAATGTGGC
YLR008C	Up	TGGTCCAGCACAATCTAACA	YLR377C	Up	ACGCACACTGTTTAGAATGC
YLR024C	Up	CATCGAAGAGCACTGCATAC	YLR385C	Up	CAACCGAGTCTAATACATGC
YLR029C	Up	GCCTCAGACATCATAATAGC	YLR390W	Up	GTGACCACTACAGTTCATGC
YLR049C	Up	TGCGTCCGCCAATCAATAAA	YLR397C	Up	CTAAGACGTGGGAATTATGC
YLR051C	Up	TGCCTGAGCAAAGGGACCAA	YLR403W	Up	AGGAACAGACCATCTACTGC
YLR054C	Up	TGACCATCCAAAGTGTCCAA	YLR441C	Up	CTCCCTAGTATTCAAATGCG
YLR060W	Up	TTGCACCACAAACGTGTCAA	YML015C	Up	GCAGCTTTCTGAATATCTGG
YLR062C	Up	TGGTCAGCAACAACGCCGAA	YML035C	Up	ATAGCTCAGGAATCTCATCC
YLR063W	Up	TGACACGGCAAAGCCTCGAA	YML038C	Up	CCAGAGGAGGAAGCATATCC
YLR064W	Up	TCTGTCGAGAAATCAGGGAA	YML049C	Up	GCTGCCATGCAATAACACGA

103

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YML076C	Up	TAGACTAGCACCATACTTGC	YNL069C	Up	ACATAGTTACAAGCGGGTGA
YML077W	Up	GGGACAAACAAATCTCTTGC	YNL075W	Up	AGAGCACTACAACCCTTGGA
YML082W	Up	CCGCTGCACAGATGTTGAAG	YNL083W	Up	ACACCAGGACAAGTTTGCGA
YML086C	Up	TAAGCCATCGGCAGACAGTC	YNL088W	Up	ACAGACCCTAAAGCGTGTCA
YML092C	Up	ATTCACAGAGCGACATAGGC	YNL100W	Up	TGGAGCTATGCCCTAGTGTT
YML093W	Up	GCACTAATGCTAATTGAGGC	YNL110C	Up	TATGGCTAGGTATGACGCTT
YML102C-A	Up	GCCTTCACATCAACAGGATA	YNL112W	Up	TCACCGTGGCGAGATAGCTT
YML105C	Up	GTCGCGCATCAAGAACACGA	YNL114C	Up	TGGACCTGTGTCAGCTCCTT
YML114C	Up	CCGCGTGCCAAAGATGCAAA	YNL117W	Up	TTGTAGCGGCTCGCACGATT
YML130C	Up	GACTATGGCATCATTGTCTG	YNL119W	Up	TCCTCTGAGTGGTGGCAATT
YMR003W	Up	GCTGTCTCCAAAGCACGAAA	YNL120C	Up	TCCCTACGTCTGATGACATT
YMR005W	Up	TTTGAGCTGATCCCAGGCTG	YNL125C	Up	TATACCCGCTGAGGCTGTGT
YMR013C	Up	CTCAGAGAGCAAGCTGGATA	YNL135C	Up	TATCAGGGCCTAGTGACTGT
YMR014W	Up	CCGTGCAATCAACTTGGATA	YNL137C	Up	TATTGACGCGGCCTCGATGT
YMR021C	Up	AATAACCGAGCATGTCATCC	YNL143C	Up	TCACGTCTGAGATTGCCGGT
YMR024W	Up	AGACACCTCGAAGATGATCC	YNL149C	Up	CAGTGAGAGTTATAGAGCCT
YMR032W	Up	CCGCAGAGAGTATAAGAGTC	YNL153C	Up	GAGACTCGCTGCATTGCCAT
YMR033W	Up	ATTAAGAAGTGCGCGGAATC	YNL162W	Up	ATTACCTCTGAGCATGGCGG
YMR049C	Up	GCTGAAGCATTCCTGAGAAG	YNL170W	Up	AGCACGAGCATCATAGATTC
YMR059W	Up	ATTATACGTCCATCCAGCGG	YNL171C	Up	GCATAGCGTATAAGCGATTC
YMR066W	Up	CCGAGCTTTCTAACTAGCGG	YNL172W	Up	AAGACTGGTGAAGCAGATTC
YMR072W	Up	AAGACATCAGGATCATGCGG	YNL177C	Up	AGCAGAGCACCAGTCTAATC
YMR098C	Up	TATTGAGCCATACGCCACAG	YNL178W	Up	TCACAGCAGTAAGCCTAATC
YMR107W	Up	TAGAGTATCGGCCATCAGTG	YNL203C	Up	GCAGTATCGAAATGCACCCA
YMR113W	Up	CAGCGATACTGAGATGTCAT	YNL218W	Up	AGACTGTATCCACTACGCTG
YMR117C	Up	CACTCTTATAGATCGTGGCT	YNL223W	Up	TTCTAAGCTGCATGTAGCGG
YMR134W	Up	CAAGTCGGAGCAATGTGATA	YNL230C	Up	GACTGCCATAGACTGCATTC
YMR168C	Up	CATCTTCCGAGACTGGAGTG	YNL231C	Up	ATCGTAGACATGGCACATTC
YMR169C	Up	TGACCTCAGCCATTATTGCG	YNL234W	Up	GGCTGTAGCATAATATCCTC
YMR207C	Up	GCACGCTAGTAATCCATATC	YNL246W	Up	AGCTCAAGCGTTCGCAAGAC
YMR224C	Up	CCGATACGTGCATATTGTGG	YNL247W	Up	GCTATGCGAAGACATTTCAC
YMR227C	Up	CCCTGACTATGATATGAGTG	YNL254C	Up	CCTGACTACCAATGCAAGGA
YMR257C	Up	CACGCACGTCGATCTATGTG	YNL282W	Up	ATGTGTATGCCACTACCGCG
YMR258C	Up	CTGATTCCACGAGGTATGTG	YNL328C	Up	TGCTCATGCCGATCTGCGAT
YMR263W	Up	ACACGGAGGTCATGCTATCT	YNL331C	Up	TCACTGTATGCAGACTGGTG
YMR280C	Up	GCGATTGGAACACATGAGAC	YNL332W	Up	AGGCATTTGTCACACTGGTG
YMR281W	Up	CCTTGATGAGAAGATGAGAC	YNR003C	Up	CCTCTTGCATAATACGGTAG
YMR284W	Up	AATAGCTCTTCAAGGCCAGC	YNR014W	Up	AACCACCCTCGATATTGTAG
YMR290C	Up	TAGCAGCCAGCATTAGGCAG	YNR015W	Up	CACTCATGGTTGGGAATTAG
YMR290W-A	Up	TACGAGTCTACAGCAGGCAG	YNR016C	Up	CGACCACTTGTCTGAATTAG
YMR297W	Up	CCATCTGGGTTAAGTGCATG	YNR017W	Up	GAGTTACTACCCACCATTAG
YMR300C	Up	CCCGCAGCTAGACTTATGTG	YNR021W	Up	ACACCGTCACCATAGATTAG
YNL011C	Up	ACGACTACCAGATTCACTAG	YNR023W	Up	GCCTCTACTTAACCGATTAG
YNL014W	Up	TACACGACTGTCTAACCTAG	YNR028W	Up	CGAGCAGGGTACATTATTAG
YNL016W	Up	ATGACGAGCCCATGACCTAG	YNR033W	Up	CCCATCCTTTGGAATCTTAG
YNL021W	Up	AGAGTCATCCCATTACCTAG	YNR035C	Up	GCCGTATCTTAACCTCTTAG
YNL042W	Up	GCACCCATCTTCATAAGTAG	YNR038W	Up	GCGGCTATCAGATTAGTTAG
YNL051W	Up	ACACGACGGCAAGCCTAGTA	YNR044W	Up	ACGGATCGGACATTGTTTAG
YNL059C	Up	TAGCGTAGCGCATCTTCGTG	YNR046W	Up	GCATCCAGTAGTGGCAAACG
YNL061W	Up	AGGATGGGCCAATCTCCCTA	YNR073C	Up	AAATAACCCGCAGTAGTCGG
YNL067W	Up	AATCACGCCCAATACGTTGA	YNR074C	Up	AACCTATGTGGACACGTCGG

104

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YOL007C	Up	GACCAGATGCACCATGTAGC	YOR206W	Up	GCTAATGTAACAGACGCTAC
YOL010W	Up	CGACAGACTACATTAAGTGC	YOR207C	Up	GCCGTCTTATCAATCAGTAC
YOL015W	Up	CCTGCATCAGCATTATGGAG	YOR217W	Up	GAGATATAGTCAACTCCACC
YOL022C	Up	GGAGACTCTTGCACATTATG	YOR222W	Up	GAGGTATGACCAGACATCCC
YOL024W	Up	CATCTCATGGCATAGAGTTG	YOR224C	Up	GCGATATTTCAAGAGATCCC
YOL029C	Up	GAGCCTTGCACCTATGCTAT	YOR227W	Up	GCGTACATATAACTACAGCC
YOL040C	Up	GCATGGTGTAAATCTCCTGA	YOR248W	Up	GATGACCTCTAACATTGTCC
YOL042W	Up	GCATCATTACAACACGCTGA	YOR250C	Up	GAGAGCCCATAATCTATTCC
YOL049W	Up	ACATGGCACGAAGCTATTAC	YOR254C	Up	GATGGCACACCACTTTAAGC
YOL052C	Up	CATAGTACGGAAGATGTAGC	YOR260W	Up	GGCCTTACCATCAATATAGC
YOL060C	Up	CATCATGCCTAACTCTGGAG	YOR261C	Up	GGCCACTTCTAACATATAGC
YOL069W	Up	CTTCTGAGAGGAGACTTATG	YOR262W	Up	GCTTAACATCACGTACTAGC
YOL110W	Up	AGAGACCACTCATTTCAGGG	YOR264W	Up	GTCACACAGCTAAGTAACGC
YOL120C	Up	AACTTCACCGGCATTGAGGG	YOR265W	Up	GCTAACTATGTTCAAGACGC
YOL123W	Up	CTCCAGATTATACGATAGGG	YOR266W	Up	GTAGATAATCCACCAGACGC
YOL127W	Up	TATACCTTCAGAGAGTAGGG	YOR269W	Up	GTGAGACATATAACCTCCGC
YOL130W	Up	CTCTGTATGAGCATTTAGGG	YOR281C	Up	GATGAGGACTAACTCCCTGC
YOL135C	Up	TATTTCAGGGCCATCACGGG	YOR282W	Up	GAACGACGGCGAATATCTGC
YOL144W	Up	CCTCGCAGTAAAGATGAGCA	YOR284W	Up	GCCGACTATACATTACTTGC
YOL147C	Up	CGCTATCATGTCATCGAATG	YOR294W	Up	AGATCACATCACTCGATAGC
YOL153C	Up	CAGGCAATCTTAGCCAAGTA	YOR318C	Up	GCAGGCTCTATCACCTATAT
YOR008C	Up	CTCTCAGGAGGAAGAATGTA	YOR320C	Up	ATTCTGAGCGGTGCCATAGT
YOR016C	Up	TATTATCAGACATCGCACGC	YOR330C	Up	GCGGCCATCACATATTATCA
YOR017W	Up	GAGAGATAGCCACATCACGC	YOR331C	Up	GCGGCATTAACAGCTTATCA
YOR031W	Up	ACTCATATTGCATGGCACTG	YOR335C	Up	GCAGTCCTACCAATCTATGA
YOR032C	Up	GTGTATCATCCCTATCACTG	YOR336W	Up	CCGCCACAGATCAATTAGTA
YOR048C	Up	GCAGCGCAGAAAGTCTCACA	YOR347C	Up	ATGAGCCAGGAACCTCTAGC
YOR056C	Up	CGAGTTACAACAGTCTCTGA	YOR361C	Up	TCATGCAGGCCAGTGCTATG
YOR057W	Up	CAGCCTGTTAAATCGTCTGA	YOR362C	Up	ACGAGTCTCTGACATCTATG
YOR063W	Up	CGCGCACATATAGGAAGACC	YOR367W	Up	ATCTCATGCGTACCGACTAT
YOR066W	Up	GCTTAGCACATACATAGACC	YPL008W	Up	TGCGACCATTGGCGATAACG
YOR110W	Up	CCACTGTATGGAAGATCATC	YPL011C	Up	TGAGAGATATACGGCCTTAG
YOR118W	Up	AGGCATCTCACATACTGTTC	YPL024W	Up	TATCCTACTACAGTCGTCAG
YOR123C	Up	TATCTGAGACGATTGCATGG	YPL025C	Up	TATTTCCAGCCAGGGTGCAG
YOR127W	Up	GCAGGCTTTGTAATGATCTG	YPL029W	Up	TAGCGGTCACCATTACGCAG
YOR128C	Up	CCTTCATTGACACTGATCTG	YPL041C	Up	TAGCACCGACTACAGTGTTC
YOR130C	Up	CTAGCTGATCGCACTATCTG	YPL046C	Up	TACGACGTGAGACCATCTTC
YOR133W	Up	GGCATATACGCTACTGGCAT	YPL047W	Up	TAGCTGGTAACAGTTGCTTC
YOR138C	Up	ATTTATGCGCGACGCCAGCT	YPL051W	Up	TACAGGGTCATAAGCGTGTC
YOR145C	Up	CTTCCCGACAGAGCAAGAGA	YPL063W	Up	TGGACGGCATAAGTCTCCTC
YOR146W	Up	CCTCCGCAGCAATTAAGAGA	YPL064C	Up	TAAGTCGAGATAATCGCCTC
YOR157C	Up	GCGACCTAGACATTATCATC	YPL065W	Up	TGAACATCCGTGAACCTATC
YOR159C	Up	CCAAGAGTGCGGAATATCTC	YPL068C	Up	TATGAGGTTCCACACCGTGC
YOR161C	Up	CCTTACGCTGAATGTATCTC	YPL070W	Up	TAAGCTAGTATAACCGCTGC
YOR168W	Up	TTGCGGCATGTGACACTCTG	YPL076W	Up	TATATTAGACCAGAGGTGGC
YOR181W	Up	CGCTGATAGTTAGCTGGATT	YPL082C	Up	TAAGAGCAGATAAGTTCGGC
YOR182C	Up	CAAGCCTGGGTCACATGATT	YPL083C	Up	TAATTGCCTCGAATCTCGGC
YOR186W	Up	TCCGGCGATCAACAGAGAGA	YPL088W	Up	TAGGTAATGACACTGAAGGC
YOR189W	Up	GCCCGAGAATCCAAGGCATA	YPL089C	Up	TATGCGCTAACAGGGTTCGC
YOR202W	Up	GCGACTATCGAACCATATAC	YPL106C	Up	CAGCTACTGGAATAGTTTGC
YOR204W	Up	GCTACAGACGAAGGCACTAC	YPL113C	Up	TAGTATATGACACCACAGCC

105

Knockout ORF	Tag type	Tag sequence	Knockout ORF	Tag type	Tag sequence
YPL122C	Up	CCGCCGACATAAGATATTGA	YPR183W	Up	CTGGAGTAATACCGACTAAG
YPL142C	Up	CTAGATACGTTCGCAGACAT	YPR194C	Up	GCCACGTCATCTAATCTAAG
YPL143W	Up	GCAGCGACTACGTTGCACAT			
YPL154C	Up	GCACTCAGATCATATCCTAG			
YPL157W	Up	GCATCTCTATAACTGGTGTC			
YPL163C	Up	AGAGCATACGCTATAACCTC			
YPL165C	Up	CCGTCATCATAACTTGTTGC			
YPL181W	Up	GAATCCATGCCAACCATTGA			
YPL189W	Up	GCCAACGTGAAATATCCGCA			
YPL204W	Up	TACTATCTGCCATGCCGGTG			
YPL207W	Up	CACTATACGTTGAGACCATG			
YPL209C	Up	CCTACGCAGAGATTATAGCG			
YPL216W	Up	TAGACTGCGCTAACTGCCAG			
YPL217C	Up	ATGCAGACTTCCAGATGGTG			
YPL218W	Up	GAGATACTCCACACGATGTC			
YPL223C	Up	AGATACTAGACATGCCACTC			
YPL227C	Up	GCTCATCTAACAGTATTGGC			
YPL228W	Up	AACAGCCTGGAAGCATTGGC			
YPL233W	Up	GCGATGCTATCAATACTCAC			
YPL235W	Up	TGAAGACTACTGCAACTCAC			
YPL238C	Up	CCGGAGCATCAATATGACTA			
YPL243W	Up	GAGATCCTGACAACCATCGA			
YPL244C	Up	GGCACATCGCAAGACATCGA			
YPL245W	Up	GGCATGACAACAGATCCGCA			
YPL249C	Up	GCTCCGTTCAAACAGATCAA			
YPR016C	Up	CCACGCAGTACAGATTGGTC			
YPR019W	Up	ACTAAGATCCCATGTTGGTC			
YPR021C	Up	GTTCGATAACTAGCCATGTC			
YPR034W	Up	AGAGACCGCACTAACGTGTC			
YPR041W	Up	CAGAAGAGTTAAGCCTTGTC			
YPR044C	Up	GACCACCCTAGATATTTGTC			
YPR053C	Up	GCAAACAGACTCGGTAATTC			
YPR063C	Up	GGTCTACCTACATAGCATTC			
YPR069C	Up	GCGTACCTAAGATCAGATTC			
YPR086W	Up	CACGGTCTTTAATAGTGGTG			
YPR099C	Up	TAGTGAGGACTCCCAGCTTC			
YPR108W	Up	GGTGATCGACACTCTCCAAG			
YPR126C	Up	GGTCTCACACTATCTTCAAG			
YPR130C	Up	GGTTCATAATGTCGCAGAAG			
YPR131C	Up	GCTATCTAAGTGGTCAGAAG			
YPR132W	Up	CTTGACCACATCCAGAGAAG			
YPR137W	Up	CACTCGTAACTATACCGAAG			
YPR144C	Up	TGAGACGCTACCACAGGAAG			
YPR148C	Up	TTACCCTAAGTGGACGGAAG			
YPR157W	Up	CTCGTTGAAGGTTACTGAAG			
YPR159W	Up	GCTATCTAACTCTCCTGAAG			
YPR165W	Up	CCTAGTACAGGAGATTGAAG			
YPR169W	Up	CCGGCCATGTATAGAATAAG			
YPR175W	Up	GCCATTACAGAGGCGATAAG			
YPR177C	Up	CCGCACCCAGTTAATATAAG			
YPR178W	Up	CGACCTGCATTCTAACTAAG			

106

Chapter 5.

Gene function prediction from congruent synthetic lethal interactions in yeast

Ping Ye[†], Brian D. Peyser[†], Xuewen Pan, Jef D. Boeke, Forrest A. Spencer[‡] and Joel S. Bader[‡]

[†] These authors contributed equally to this work.

[‡] Corresponding authors.

Abstract

We predicted gene function using synthetic lethal genetic interactions between null alleles in *Saccharomyces cerevisiae*. Phenotypic and protein interaction data indicate that synthetic lethal gene pairs function in parallel or compensating pathways. Congruent gene pairs, defined as sharing synthetic lethal partners, are in single pathway branches. We predicted benomyl sensitivity and nuclear migration defects using congruence; these phenotypes were uncorrelated with direct synthetic lethality. We also predicted *YLL049W* as a new member of the dynein-dynactin pathway and provided new supporting experimental evidence. We performed synthetic lethal screens of the parallel mitotic exit network (MEN) and Cdc14 early anaphase release pathways required for late cell cycle. Synthetic lethal interactions bridged genes in these pathways, and high congruence linked genes within each pathway. Synthetic lethal interactions between MEN and all components of the Sin3/Rpd3 histone deacetylase revealed a novel function for Sin3/Rpd3 in promoting mitotic exit in parallel to MEN. These *in silico* methods can predict phenotypes and gene functions and are applicable to genomic synthetic lethality screens in yeast and analogous RNA interference screens in metazoans.

Introduction

The robustness of a biological network to defects can be probed by synthetic lethality, which reveals that a cell survives individual gene deletions, but cannot survive deletion of specific gene pairs. Synthetic lethal interactions have been rationalized with two hypotheses: (i) two genes in a single linear pathway can show synthetic lethality; (ii)

108

synthetic lethal genes act in parallel or compensating pathways (Tucker and Fields 2003). These two hypotheses predict distinctly different patterns of synthetic lethality: enrichment of interactions within single pathways versus depletion of interactions within pathways and enrichment between pathways. These two hypotheses also make different predictions for the nonlethal phenotypes of the underlying single gene deletions: a shared phenotype for genes in a single pathway, or possibly differing phenotypes for genes in parallel pathways.

Hypothesis (i) is possible only when alleles are hypomorphic but not complete loss-of-function mutants: each mutation reduces flux partially, but the combined reduction from two mutations leads to lethality. Hypothesis (i) does not apply to synthetic lethality between null alleles, with complete loss of function. Hypothesis (ii) is expected in this case, with each null mutation knocking out one of the two parallel pathways that sustain normal growth. In this view, an essential protein complex that retains function when single nonessential subunits are deleted (but not multiple subunits simultaneously) is formally represented by multiple pathways, one for each functional stoichiometry, connected in parallel.

Data sets to test these rationales are arising from high-throughput synthetic lethality screens accomplished in *Saccharomyces cerevisiae* using synthetic genetic array (SGA) and synthetic lethality analysis on microarrays (SLAM). These screens test a deletion of interest (query gene) against all possible viable yeast single-deletion strains (target genes) (Tong et al. 2001; Ooi et al. 2003; Pan et al. 2004). As human disease susceptibility may encompass gene mutations in multiple pathways, synthetic lethality is relevant to human

disease processes (Tucker and Fields, 2003).

We focus on the subset of genetic interactions restricted to synthetic lethal interactions and synthetic fitness (slow growth) defects between *null alleles*. These interactions are easier to interpret than more general genetic interactions (enhancer, suppressor screens) or other types of mutant alleles (e.g., hypomorphs of essential genes). Null mutants constructed by the International Yeast Gene Deletion Consortium represent the vast majority currently under study by the yeast community (Giaever et al. 2002). For brevity, we use the term synthetic lethal to include both the lethal and reduced fitness phenotypes.

Synthetic lethal interactions have been used to predict that interaction partners share function in the same pathway (Tong et al. 2001, 2004; Wong et al. 2004). Here, we emphasize the alternative hypothesis suggested above, that synthetic lethal interactions bridge parallel pathways, which are in a sense orthogonal to direct synthetic lethal interactions (Figure 1A). This concept is formalized computationally as follows. Pathway membership is inferred using the hypergeometric P-value for a shared pattern of interaction partners, which we abbreviate as the congruence score (Figure 1B).We present evidence that functional associations inferred from the congruence score are stronger than associations between the synthetic lethal interaction partners themselves. Two types of functional associations are explored: biochemical participation in protein complexes, through joint analysis of synthetic lethal interactions (Tong et al. 2004) with protein complex data (Gavin et al. 2002; Ho et al. 2002; Mewes et al. 2004) (see Supplementary information, Supplementary Figures S1 and S2); and phenotypes of the

110

underlying single gene deletion mutants, including nuclear migration and drug sensitivity. The nuclear migration assay and the physical interaction detected between Jnm1p and Yll049wp confirm our prediction that the previously uncharacterized yeast gene *YLL049W* is a new member of the dynein–dynactin pathway.

Results

Congruent genes function in the same pathway

As has been noted previously, only ~1% of synthetic lethal interactions occur between genes whose products reside in a single protein complex (Tong et al. 2001). While, as pointed out by the authors of that paper, this is a greater fraction than would be expected by chance, it is clear that the vast majority of synthetic lethal interactions are not explained by common protein complex membership and we would argue that this 1% represents the exception and not the rule. The parallel pathway model suggests that genes sharing synthetic lethal interaction partners may function in a single pathway, and their gene products should have an increased probability to reside in a single protein complex.

The raw number of shared genetic interaction partners has been used previously to rank the probability of a physical interaction between the corresponding gene products (Tong et al. 2004). Here, we instead use the hypergeometric *P*-value for the number of shared neighbors, which accounts for the number of interaction partners of each gene (Figure 1B). To convert this value to a convenient scale, we define the congruence score as the negative $\log_{10}$ of the *P*-value; related measures have been used to analyze protein interaction networks (Goldberg and Roth 2003; Schlitt et al. 2003) and multiple

characters from single RNA interference (RNAi) screens (Gunsalus et al. 2004). The

congruence score has the benefit of providing a natural significance threshold

incorporating the size of the network. The performance of a predictive method can be

visualized by plotting the number of true positives versus the number of false positives as

a function of the number of predictions made, known as a receiver operating

characteristic (ROC) curve. Based on the area under the ROC curve, the performance of

congruence score method is superior to counting the number of shared partners in

predicting protein complex membership in the stringent regime (Supplementary

Figure S3 and Supplementary Table S1).

We separated the synthetic lethal interaction data into 'query' and 'target' sets,

based on whether each gene node represents a non-essential query gene (126 are included

in the published data) or a target gene (982 of which are synthetic lethal partners of at

least one query).We calculate congruence scores for each pair of target genes

(Supplementary Figure S4).

The fraction of target gene pairs in the same protein complex (Gavin et al. 2002; Ho

et al. 2002) increases with congruence score, rising to 100% at the highest values

(Figure 2A). Analysis using the MIPS database of curated complexes (Mewes et al. 2004)

yields similar results (Supplementary Figure S5). Even for the smallest non-zero

congruence scores, the observed fractions of pairs within the same complex are greater

than expected by chance ($P < 0.005$). Gene products of pairs with congruence score ≥ 5

have a higher probability of protein complex co-residence than products of synthetic

lethal interaction partners. Moreover, using synthetic lethal interactions to predict

complex co-residence shows higher false positive rate ([false positives]/[false positives+true negatives]) and higher false discovery rate ([false positives]/[false positives+true positives]) than using congruence score (Supplementary Figure S3).

Functional associations, determined by extracting Gene Ontology (GO) (Ashburner et al., 2000) annotations and calculating correlations based on the depth of the deepest parent term (see Materials and methods), are greater for congruent genes than for synthetic lethal pairs. Biological Process and Cellular Component correlations increase with congruence score and are greater than the similarity between direct genetic interaction partners (Figure 2B). As is typically the case, the GO Molecular Function annotations have smaller correlation as they refer to molecular, rather than biological., roles. For congruence scores ≥ 7, ≥ 10, and ≥ 6, respectively, the GO process, function, and component correlations for congruent gene pairs are significantly higher than the corresponding correlations for the raw synthetic lethal pairs (0.25, 0.05, and 0.31), respectively ($P < 0.05$). Calculations based on semantic similarity of GO terms (Lord et al. 2003) show even stronger performance of the congruence score relative to synthetic lethality (Supplementary Figure S12).

In summary, a congruence interaction with score ≥ 10 provides a tighter functional relationship than synthetic lethality, consistent with our interpretation of single versus parallel pathways. Although individual synthetic lethal gene pairs may share synthetic lethal partners (as observed by Tong et al. 2004), high congruence score typically excludes direct synthetic lethal interaction, in agreement with our model (Figure 2C). When congruence score is greater than or equal to 14, the binomial P-value for observed

number of synthetic lethal interactions becomes insignificant given the overall frequency of synthetic lethal interactions observed in the entire congruence data set ($P > 0.05$).

A network generated by setting a threshold congruence value ≥ 10 recapitulates known functional associations and suggests novel associations (Figure 2D). Sets of genes known to function within the same pathway tend to cluster together. As expected, the congruence links overlap known protein interactions, whereas synthetic lethal links do not. For example, a prefoldin complex gene cluster inferred from congruence links (*PAC10*, *GIM3*, *GIM4*, *GIM5*, and *YKE2*) corresponds to the *PAC10* complex shown in Supplementary Figure S1B.

In some cases where proteins encoded by genes with congruence links were not detected within the same protein complex by high-throughput studies (Gavin et al. 2002; Ho et al. 2002), other experiments have indicated physical interactions. *SWR1*, *SWC1*, *VPS71*, *VPS72*, *SIF2*, and *ARP6* encode subunits of *SWR1* chromatin remodeling complex catalyzing exchange of histone H2A with histone variant Htz1p (Mizuguchi et al. 2004). Genes in a highly connected congruence cluster may function in the same pathway through transient physical interactions, or they may participate in a pathway as separate physical entities. For example, Cin1p, Cin2p, and Pac2p are all tubulin folding factors that function in a pathway leading to microtubule stability (Hoyt et al. 1997). Physical interaction between Pac2p and Cin1p has been reported (Fleming et al. 2000). Cin8p is a kinesin motor protein involved in mitotic spindle assembly and chromosome segregation, and interacts with microtubules (Gheber et al. 1999). Possibilities include that Cin1p, Cin2p, Pac2p, and Cin8p interact transiently during mitosis, or that they

114

influence the same molecular environment independently. For example, activities of Cin1p, Cin2p, and Pac2p might generate an optimal microtubule substrate for Cin8p.

The largest connected component in Figure 2D includes known members of the dynein–dynactin spindle orientation pathway (*ARP1*, *NUM1*, *DYN1*, *PAC11*, *PAC1*, *DYN2*, *JNM1*, *YMR299C*, and *NIP100*) and corresponds to a group observed previously using clustering (Tong et al. 2004). The dynactin protein complex (Arp1p, Jnm1p, and Nip100p) defined by biochemical studies is required for proper spindle orientation and chromosome partitioning to daughter cells during anaphase (Kahana et al. 1998). Additional reported protein–protein interactions in this congruence cluster include Jnm1p–Yll049wp, Nip100p–Pac11p, Pac11p–Dyn2p, and Pac11p–Num1p (Uetz et al. 2000; Farkasovsky and Kuntzel 2001; Ito et al. 2001). We predict *YLL049W* as a new component of the dynein–dynactin spindle orientation pathway, which is consistent with previous observation (Tong et al. 2004). We have experimentally validated the functional prediction of *YLL049W* by showing that its null mutant allele exhibits a nuclear migration defect similar to dynactin component *JNM1*. Furthermore, we have successfully detected a physical interaction between Jnm1p and Yll049wp using a directed two-hybrid test. Both experiments will be described in detail in the next section. The second uncharacterized open reading frame (ORF), *YDR149C*, is also congruent to dynein–dynactin components. Its ORF overlaps the beginning of its neighbor *NUM1*, and we suggest that the *ydr149cΔ* phenotype is in fact due to concomitant mutation of *NUM1*.

Congruence scores predict pathway components and quantitative phenotypes

Distinct lesions to a single pathway branch should result in similar systems-level

perturbations. We reasoned that similarity of a numeric phenotype of a deletion mutant should be better predicted by congruence score than by a direct synthetic lethal interaction.

We investigated the ability of the congruence score to predict the penetrance of nuclear migration defects in a population of mutant cells. Mutations in the dynein–dynactin spindle orientation pathway are known to increase the nuclear migration defect rate. We selected six genes in the pathway as landmarks (*DYN1*, *ARP1*, *DYN2*, *JNM1*, *NUM1*, and *NIP100*) and then measured the defect rate at 13 ºC for 59 mutants of genes with congruence score ≥ 4 to at least one of the landmarks (Supplementary Figure S6 and Supplementary Table S2). To summarize the relationship between phenotype and congruence score, each mutant's migration defect (% abnormal) was plotted as a function of congruence scores to landmark genes (Figure 3A). The average congruence score is highly correlated with the defect rate (Spearman correlation coefficient = 0.51, two-sided $P = 3.9 \times 10^{-5}$). Additionally, at or above congruence score of 10, all mutants exhibit moderate to severe nuclear migration defects (14–80% abnormal cells).

Among the mutants found to exhibit a nuclear migration defect was one representing the unstudied gene *YLL049W* (Supplementary Table S2). Further analysis of the *yll049w* mutant showed that the observed defects are temperature-dependent, similar to *jnm1* mutants, whereas a mutant for the Kinesin-related *KIP2* gene displayed temperature-independent defects (Supplementary Table S3). Notably, the *JNM1*–*YLL049W* congruence score (15.2) is higher than the *JNM1*–*KIP2* congruence score (10.8), consistent with more similar phenotypes.

116

It is evident from this analysis that uncharacterized ORF *YLL049W* is required for robust nuclear migration. High-throughput yeast two-hybrid results suggested a protein–protein interaction between Yll049wp and dynactin subunit Jnm1p (Ito et al. 2001). We have experimentally confirmed this physical interaction between Yll049wp and Jnm1p using a different two-hybrid system (Supplementary Figure S7). These results provide supporting evidence for interaction between the two proteins, but do not address whether the association is stable, transient, or bridged by other proteins. The dynein–dynactin pathway for nuclear positioning includes many protein components that are not dynein or dynactin complex members, whose contributions influence microtubule dynamics, the formation of a capture site on the cell cortex, and proteins that regulate spatial and temporal steps in the determination of nuclear orientation and migration during the cell cycle (Sheeman et al. 2003; Knaus et al. 2005; Li et al. 2005). Kip2p acts to ensure nuclear positioning within the dynein–dynactin pathway (Miller et al. 1998) by transporting dynein to the microtubule plus ends (Lee et al. 2003; Carvalho et al. 2004).

Our data indicate that *YLL049W* is a previously unknown component of the dynein–dynactin spindle orientation pathway and suggest that it might be a subunit of yeast dynactin. Elucidation of the specific molecular function of *YLL049W* will require further study.

To test the general application of using congruence score as phenotype predictor, we chose sensitivity to benomyl, a microtubule-depolymerizing agent, as our second phenotype assay for deletion mutants. The microtubule biogenesis gene *CIN1* (Hoyt et al. 1990) was selected as the benomyl-sensitive landmark. Null mutants of 31 genes with

congruence scores ≥ 4 for *CIN1* were tested for growth defects on medium containing

5 mg/ml of benomyl at 25 ºC (Supplementary Table S4). With increasing congruence

score cutoff, the fraction of benomyl-sensitive null mutants rises to 1 (Figure 3B). We

again observed significant correlation between the congruence score and the fraction of

benomyl-sensitive mutants (Spearman correlation coefficient = 0.49, two-sided

$P = 0.006$).

To validate the hypothesis that congruence interaction inferred from synthetic

lethality indicates a closer functional association between genes than direct synthetic

lethality, we selected landmarks of seven benomyl-sensitive mutant strains (*cin1Δ*,

yml094c-aΔ, pac10Δ, pfd1Δ, gim3Δ, tub3Δ, and *gim5Δ*) from the top list of 451

candidate benomyl-sensitive mutant strains from a recent high-throughput genetic screen

(Pan et al. 2004).We then ranked genes based on their average congruence score with

seven landmarks (Supplementary Table S5). As a test of the competing hypothesis that

synthetic lethal interactions themselves indicate direct functional associations, we also

ranked genes by the raw number of synthetic lethal interactions with seven landmarks

(Supplementary Table S6). The congruence score and the raw number of interactions

were then tested for correlation with benomyl LD_{50}, the dose that is lethal to at least 50%

of the cells, equivalent to control/experimental hybridization signal ratio ≥ 2 used as

threshold by Pan et al. (2004). The congruence score is significantly correlated with LD_{50}

(Spearman correlation coefficient = −0.17, two-side $P = 0.04$), but the number of

synthetic lethal links is not (Spearman correlation coefficient = 0.06, two-side $P = 0.22$)

(Figure 3C and D). These results support the idea that genetic congruence correlates

better with a given phenotype than direct synthetic lethal interaction and indicate that

congruence is a superior measure for predicting certain phenotypes.

All genes having high congruence scores with landmarks are involved in direct

microtubule biogenesis. For example, *PAC10*, *YKE2*, *GIM3*, *GIM4*, and *GIM5* all belong

to the prefoldin complex that acts to deliver unfolded proteins to cytosolic chaperonin

(Geissler et al. 1998; Vainberg et al. 1998). On the other hand, we noticed that some

genes with multiple synthetic lethal interaction links with landmarks tend to function in a

distinct pathway from microtubule biogenesis. For example, *SWC1* and *ARP6* are

subunits of *SWR1* chromatin remodeling complex catalyzing exchange of histone H2A

with histone *HTZ1* (Mizuguchi et al. 2004).

Physical co-residence predicts genetic congruence

Because increasing congruence score is related to protein complex co-residence, we

predicted that genes encoding proteins known to co-reside in a complex would have

similar synthetic lethal interaction profiles. We verified this hypothesis using *PFD1* as a

dSLAM (diploid-based synthetic lethality analysis on microarrays) query; the remaining

prefoldin complex members have been characterized as queries in the SGA study. We

identified 33 *PFD1* synthetic lethal partners (Supplementary Table S7). High congruence

values between *PFD1* and other prefoldin components, *GIM3*, *GIM4*, *GIM5*, *PAC10*, and

YKE2, equal to 14, 14, 9, 15, and 16, demonstrate the overlap between congruence links

and protein complex membership (Supplementary Table S8). The five prefoldin members

used as query genes in SGA exhibit much more significant overlap among themselves

(congruence scores in the range of 23–67) than to *PFD1*. However, this may arise from

119

systematic biases between the SGA and dSLAM methods rather than a biological

distinction for *PFD1*. Additionally, 13 of 33 *PFD1* synthetic lethal partners map to

reported protein complexes (Supplementary Table S7). Notably, none of the 33 *PFD1*

synthetic lethal partners is a prefoldin component. This supports the hypothesis that

physical and synthetic lethal interactions are generally orthogonal.

Synthetic lethal interactions predict parallel pathways

We further tested the hypothesis that synthetic lethal interactions between null

alleles define parallel pathways, by performing dSLAM screens of genes required for

mitotic exit. Two parallel pathways, the Cdc14 early anaphase release (FEAR) and the

mitotic exit network (MEN), are required for release of the essential protein phosphatase

Cdc14p from nucleolus during yeast cell cycle (Stegmeier et al. 2002). Components of

the FEAR network include *SLK19* and *SPO12*, whereas those of MEN include *LTE1* and

CLA4. Double mutant cells of these two pathways fail in Cdc14p release from the

nucleolus and arrest in telophase with a large-budded morphology.

To test the parallel pathway model, we performed dSLAM experiments using

SLK19, *SPO12*, and *LTE1* as queries; *CLA4* was previously used as a query in the SGA

study (Tong et al. 2004) (Supplementary Table S9). We re-identified known synthetic

lethality interactions between the FEAR and MEN pathways (Stegmeier et al. 2002;

Goehring et al. 2003). High congruence was observed between *SLK19* and *SPO12*, and

between *LTE1* and *CLA4*, but not across FEAR/MEN pathways (Supplementary

Table S10). In addition, our genome-wide screens discovered synthetic lethal interactions

between *LTE1* and *SIN3*, *RPD3*, *PHO23*, and *SAP30*, the components of the Sin3/Rpd3

histone deacetylase complex (Loewith et al. 2001) (Figure 4A and Supplementary

Table S9). Although most were initially not identified in the previous study (Tong et al.

2004), we also observed synthetic lethal interactions between *CLA4* and all four

components of the Sin3/Rpd3 complex (data not shown). These interactions were specific

to MEN because synthetic lethality was not observed between the Sin3/Rpd3 histone

deacetylase components and FEAR network components in either dSLAM or individual

assays. These results led us to predict that the Sin3/Rpd3 histone deacetylase might play

an important role during mitotic exit when the MEN pathway is mutated. In support of

this, cells of the double mutants, *lte1Δ rpd3Δ, lte1Δ sin3Δ, lte1Δ sap30Δ, lte1Δ pho23Δ,*

were unable to exit mitosis, and arrested with a dumbbell-shaped morphology typical of a

mitotic exit defect. Furthermore, the viability of these double mutants was restored when

TAB1-6, a dominant allele of *CDC14* that binds weakly to the negative regulator

Cfi1p/Net1p (Shou et al. 2001), but not the wild-type *CDC14* was expressed (Figure 4B).

Interestingly, this *TAB1-6* allele also suppressed the lethality of an *lte1Δ slk19Δ* double

mutant (Figure 4B). Thus the Sin3/Rpd3 histone deacetylase likely acts in parallel with

MEN in promoting exit from mitosis.

Discussion

Synthetic lethal interaction provides evidence for compensating gene function. This

compensation has been rationalized as buffering within a single pathway, or buffering

between two parallel or compensating pathways (Tong et al. 2001, 2004; Wong et al.

2004). We find that the parallel pathway model permits successful inference of protein

complex membership from synthetic lethal data. The parallel pathway model, but not the

single pathway model, yields successful predictions for phenotypes including nuclear migration defect rates and drug sensitivity. The parallel pathway model is also consistent with known pathways comprising genes identified in synthetic lethal screens. The model motivated our confirmation of *YLL049W* as participating in the dynein–dynactin nuclear migration pathway by phenotypic analysis, permitted identification of benomyl-sensitive strains based on congruence to landmark genes, and yielded a novel prediction of Sin3/Rpd3 histone deacetylase as a new module for mitotic exit that acts in parallel with MEN.

Using a different analysis strategy, Kelley and Ideker (2005) recently reported that synthetic lethal interactions are typically 'between pathway', whereas 'within-pathway' interactions occur infrequently. For their purposes, all subsets of proteins that are densely connected by physical interactions in non-mutant cells were considered 'within pathway'. If a pathway is defined strictly by its components, however, the view that null allele synthetic lethality must always occur between parallel pathways can be enforced, precluding 'within-pathway' explanations. In such a view, members of a protein complex that functions in the absence of either of two subunits, but not both, would participate in three parallel pathways: one that includes all possible components, and one for each 'incomplete' complex (all of which might function in non-mutant cells). More generally, methods that summarize synthetic lethal relationships are often more useful than raw synthetic lethal pairs.

This recent analysis also predicted that Yll049wp associates with dynactin during spindle orientation (Kelley and Ideker 2005), consistent with our observation from

congruence analysis that *YLL049W* is functionally related to dynein–dynactin pathway. Our characterization includes experimental validations that support the prediction, and provides evidence from congruence score and detailed phenotype that the function of *YLL049W* is more similar to *JNM1* than *KIP2*. Confirmation of a physical interaction between *YLL049W* and *JNM1* further suggests that the prediction will be useful in future detailed analysis of the molecular role of *YLL049W*.

The congruence score metric compares favorably with other methods for inferring functional associations from synthetic lethal data. First, it produces stronger inference of gene function than the underlying direct genetic interactions. For example, direct interactions are unable to predict benomyl sensitivity, whereas congruence is a strong predictor of similar sensitivity. Second, the congruence metric naturally provides a *P*-value and can give improved performance relative to the raw count of the number of shared interaction partners. Finally, the *P*-values provided by the congruence score can provide an advantage over methods such as hierarchical clustering, which continue to depend on visual inspection of clusters and definition of cluster boundaries.

The quantitative characteristic of each congruent pair interaction can be used to consider interactions above a given threshold, allowing experimentalists to consider which network features reflect the most significant evidence in the data set, and to include less significant observations to be evaluated when desired. Importantly, a congruence summary at any significance level quantitatively relates genes according to their functional similarity by interaction profiles, not individual synthetic lethal pairings. To identify congruent gene pairs with greater or lesser significance, the interaction

linkages can be annotated, or the map can be redrawn at differing congruence cutoff scores. For example, Supplementary Figure S8, Figure 2D, and Supplementary Figure S9 are all target gene congruence network by setting congruence score ≥ 8, ≥ 10, and ≥ 15, respectively. This aspect of network analysis will become increasingly important as the information summarized within it grows. Some biologically important relationships may inherently be present in the genetic congruence network only at relatively low significance overall. These can be viewed by extracting a local network containing first-degree congruence relationships in much the same way as the current large-scale interaction network is commonly viewed in subsections (Tong et al. 2001, 2004; Ooi et al. 2003).

A possible limitation of our analysis is the low coverage of the synthetic lethal network, with only ~2% screened by high-throughput methods using query genes selected on the basis of specific biological themes (Tong et al. 2004). To assess the sensitivity of our analysis to missing data, and also to possible false positives, we repeated our analysis with data sets modified to contain up to 30% false positives (random interactions added to the data) and 30% false negatives (observed interactions removed from the data) (Supplementary Figure S10). Note that the false-positive rate is quite low for the SGA data owing to confirmation by tetrad or random spore analysis; false negatives are estimated in the range of 17–41% (Tong et al. 2004). Although the congruence scores shift to lower values, the overall performance is similar to using the original data set (compare Figure 2 and Supplementary Figure S10). These observations suggest that the congruence score method is robust to noisy and incomplete data.

Continuing genetic interaction screens will generate increasing volumes of data. A critical challenge is to develop computational approach to integrating these data and eventually understanding gene function. Several hurdles will need to be surmounted. Essential genes are missing from the synthetic lethal network, although they may be probed eventually using non-null mutant alleles. Certain higher-order redundancy processes may also require more than two-gene deletion to be observed. The most promising approach to ease the limitations may be to combine different types of networks for improved inference. We have performed joint analysis on genetic network and physical network to argue that the correct functional links between genes should be orthogonal to the synthetic lethal interaction (see Supplementary information). Future studies by combining other types of heterogeneous network data, such as gene expression and phylogenetic information, will certainly improve our inference of biological systems.

This work in budding yeast, made possible by the development of the comprehensive deletion collection, massively parallel phenotyping techniques, and quantitative analysis of synthetic lethal interaction data within a statistical framework, will create a template for testing and improving our understanding of biological buffering and genetic robustness in many systems as researchers gather similar information data sets from other organisms. Genome-wide synthetic lethality screens using RNAi are becoming available in other organisms (van Haaften et al. 2004) and may eventually allow analysis similar to the one we have performed in yeast. Full-genome RNAi screens have been conducted for *Caenorhabditis elegans* and *Drosophila melanogaster* (Kamath et al. 2003; Boutros et al. 2004), and genome-wide screens in other metazoans are in

progress. In instances where RNAi knockdown is complete, the congruence score method should provide a quantitative metric for shared gene function through calculating the probability of a gene pair sharing phenotypic defects in the RNAi screens. Therefore, the methodology we have applied to predict gene functions from yeast genomic synthetic lethality can be certainly extended to analogous RNAi screens for the discovery of novel gene tasks in higher organisms.

Materials and methods

Data sources

Synthetic lethal interactions, including lethal and sick phenotypes, were derived from SGA analysis in budding yeast, *S. cerevisiae* (Tong et al. 2004). We removed six essential query genes from the original 132-query gene network, including *MYO2*, *SCC1*, *CDC2*, *CDC7*, *CDC42*, and *CDC45*. The intermediate (viable) phenotypes exhibited by conditional alleles of essential genes may include loss-of-function, unregulated function, and gain-of-function aspects. In contrast, null alleles of non-essential genes are by definition solely loss-of-function mutations. We ascertained that our results and conclusions do not change when these six essential genes are included in the analysis. Yeast protein complex data were collected from two high-throughput studies, TAP and HMS-PCI, both using approaches of affinity purification of tagged bait protein to pull down complexes followed by mass spectrometry analysis (Gavin et al. 2002; Ho et al. 2002). Protein complexes that contain two or more non-essential gene encoded proteins were used (353 complexes from TAP and 427 complexes from HMS-PCI). We defined a protein complex to include the bait protein and all prey proteins detected by the bait.

126

Similar analysis was also performed using curated MIPS protein complex data set ('complexcat.scheme', June 12, 2003, 145 complexes with two or more non-essential gene encoded proteins) (Mewes et al. 2004) and results are provided in the Supplement. Pairwise protein interactions in *S. cerevisiae* derived from high-throughput yeast two-hybrid assays (Uetz et al. 2000; Ito et al. 2001) were also analyzed and found to support our conclusions (results not shown).

Randomization of synthetic lethal interactions

Synthetic lethal interactions from SGA were reported as a pair of genes directed from the query gene to the target gene. A randomized network was generated by keeping the query gene list unchanged, randomly picking one of the 982 target genes identified in the SGA screen according to the probability of each target gene shown in the interaction list with replacement, and matching it to the query gene. Duplicate query–target pairs and self-interaction pairs are rejected during randomization. Results depict the average over 10 randomizations.

Probability of congruence and congruence score

We separated the SGA interaction data into query and target sets, based on whether each gene node represents a non-essential query gene (126 are included in the published data) or a target gene (982 of which are synthetic lethal partners by at least one query). We depict results for the target genes, as the number of primary nodes is much larger and should, in principle, include the query genes.

The probability of a gene pair sharing at least k synthetic lethal interaction partners

was derived from the hypergeometric distribution:

$$p(x \geq k) = \sum_{x=k}^{min(m,n)} \frac{C(m,x) \cdot C(t-m, n-x)}{C(t,n)} \quad ,$$

in which $C(j,k)$ is the combinatorial factor $j!/k!(j-k)!$, m is the number of synthetic lethal

interaction partners for gene 1, n is the number of synthetic lethal interaction partners for

gene 2, and t is the total number of query genes (126 genes) if calculation is for a target

pair or the total number of target genes (4700 genes) if calculation is for a query pair. The

congruence score is $-\log_{10}[p(x \geq k_{obs})]$. High-scoring pairs from query genes reveal

similar patterns as target genes (Supplementary Figure S11) from the data set of Tong et

al. (2004).

To correct for multiple testing of target pairs, we estimate that a final *P*-value of

0.01 requires a per-link *P*-value of ~$0.01/982^2$, or 10^{-8}, corresponding to a congruence

score of 8 or more. For illustrative purposes, we selected a more stringent threshold of 10

(Figure 2D). At this significance, the congruence network contains only 68 nodes with

138 first-degree interactions, summarizing relationships among 1184 synthetic lethal

pairs overall.

Network visualization

Network figures were created using *Cytoscape 1.1* (Shannon et al. 2003).

GO annotation correlation

GO is held as a directed acyclic graph (DAG) to describe attributes of gene products

in three ontologies—biological process, molecular function, and cellular component

128

(Ashburner et al. 2000). To calculate the GO term similarity between a pair of genes, depths of different subbranches of the GO DAG have been recorded for each gene. Here, we assume that all of the links in the GO DAG are of equal weight. Then, the deepest depth in the GO DAG at which the pair of genes share an annotation was found and defined as depth d. Gene pairs with genes without annotation were discarded. The maximal depth Max(depths) and minimal depth Min(depths) for all genes in the synthetic lethal data set were calculated for each of three ontologies. The GO annotation correlation for a pair of congruent genes with depth d was defined by $(d–$Min(depths))/(Max(depths)–Min(depths)). For example, the maximal depth is equal to 17 and the minimal depth is equal to 1 for biological process ontology. The deepest depth for shared annotation of gene pair *JNM1* and *KIP2* is 11. Thus, the GO annotation correlation for *JNM1* and *KIP2* for biological process is calculated as $(11–1)/(17–1) = 0.63$. This is similar to the GO depth correlation in a previous study of Drosophila physical interactions (Giot et al. 2003), except that the previous study normalized the depth correlation to fall in the range 0–1. This method differs from the semantic similarity method (Lord et al. 2003) in two ways: (1) it weights GO terms by depth, whereas semantic similarity weights terms by frequency; (2) it uses the depth of the deepest annotated term, whereas semantic similarity averages over annotations. Results from the two methods are consistent (Supplementary Figure S12).

Noise robustness analysis

To account for 17–41% false negatives in the SGA data set, we randomly removed 30% of interactions from the original data assuming reported interactions are all correct.

129

To account for potential false positives (although SGA data set contains very few false positives as every interaction has been individually confirmed), we randomly replaced 30% of original interactions with random interactions. These two data sets containing false negatives and false positives, respectively, were used to repeat the congruence analysis, and this process was repeated 10 times (Supplementary Figure S10).

Experimental validation and discovery of gene function required for nuclear migration by highly significant congruence score

Null mutants of 59 genes with congruence scores greater than or equal to 4 for six landmark genes (*NUM1*, *DYN1*, *DYN2*, *ARP1*, *JNM1*, or *NIP100*) were tested for nuclear migration defects at 13 ºC. Deletion mutants were grown in YPD at 30 ºC until low-log phase and then cultures were shifted to 13 ºC for 24 h. Formaldehyde was added to 3.7% and cells were incubated at room temperature overnight. Cells were washed in 1 M sorbitol/50 mM potassium phosphate pH 7.5 (SK), permeabilized in SK + 3.7% formaldehyde + 0.5% Triton X-100 for 7 min, washed in SK, and then stained in SK + DAPI (100 ng/ml). Cells were examined under a fluorescence microscope, and 50 or 100 single large budded cells were scored for nuclear morphology. Normal cells had one DAPI mass at or through the bud neck or two DAPI masses, one in each cell body.

Experimental validation of gene function required for benomyl resistance by highly significant congruence score

Null mutants of 31 genes with congruence scores greater than or equal to 4 for *CIN1* were tested for growth defects on media containing low concentrations of benomyl at 25 ºC. Deletion mutants were grown on YPD agar, equal amounts of yeast (by OD_{600})

were suspended in water in a 96-well plate and five-fold dilutions were performed. A 96-pin device was used to transfer yeast from each well to a YPD agar plate containing benomyl (5 µg/ml in DMSO) and to YPD agar with DMSO only. Plates were incubated at 25 ºC for 3 days and scored for growth defects on benomyl versus DMSO alone.

Experimental validation of genetic congruence from physical co-residence

dSLAM was performed using *PFD1* as query gene and a pool of ~6000 heterozygous diploid knockout strains. The detailed method is described elsewhere (Pan et al. 2004). Briefly, the heterozygous deletion collection was transformed with a *PFD1* knockout construct as a pool, sporulated, and haploid double mutants were selected. Knockout-specific barcode tags were amplified with Cy3-labeled primers and hybridized to a microarray with Cy5-labeled control tags from haploid single mutants. Mutants were scored as positive only if both UPTAG and DNTAG had ratios greater than 2.0.

Experimental validation of parallel pathways predicted by synthetic lethal interactions

dSLAM was performed using *LTE1*, *SPO12*, and *SLK19* as query genes. The procedure is same as *PFD1* experiment described above. The data presented are the results of individual confirmation by random spore analysis or tetrad analysis.

Experimental validation of physical interaction predicted by congruence scores with dynein–dynactin landmark genes

Yeast two-hybrid experiments were performed using activation and binding domain vectors pOAD (*LEU2*-marked) and pOBD-2 (*TRP1*-marked), respectively, and yeast

strains PJ69-4a and PJ69-4alpha (James et al. 1996). Materials were kindly provided by Stanley Fields, Yeast Resource Center. GAL4-binding domain fusions were transformed into PJ69-4alpha and GAL4-activation domain fusions were transformed into PJ69-4a. The two strains were mated and diploids were selected on SC −Leu −Trp. The resulting diploids were plated on SC −Ade −His media in two dilutions (2 µl of 0.1 OD_{600}/ml and 0.02 OD_{600}/ml) at 30 ºC. Growth at 4 days demonstrated a strong physical interaction between Jnm1p and Yll049wp (Supplementary Figure S7).

The constructs used were JNM1-BD, JNM1 fusion with GAL4-binding domain; YLL049W-AD, YLL049W fusion with GAL4-activation domain; BD, binding domain alone; AD, activation domain alone.

Two independent JNM1-BD and two independent YLL049W-AD transformants supported growth when appropriately combined. YLL049W-BD + AD alone resulted in growth owing to self-activation and was therefore not informative (data not shown).

The plasmids and strains used for this study are distinct from those used by Ito et al. (2001), who reported high-throughput yeast two-hybrid interaction between *JNM1* and *YLL049W*.

Genotypes

PJ69-4a: *MAT*a *trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*

PJ69-4α: *MAT*α *trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*

132

We attempted further confirmation of the physical interaction between Yll049wp and Jnm1p with both co-immunoprecipitation (co-IP) and GST pull-down experiments. We first attempted to make yeast strains expressing fusion proteins (Yll029w-3HA, Yll049w-13Myc, Jnm1-3HA, and Jnm1-13Myc) by genomic integration using the Pringle cassettes that confer G418 resistance (Longtine et al. 1998). For all four cases, multiple G418-resistant integrants were selected and confirmed by PCR diagnosis. In each case, yeast extracts were prepared from two representative candidate clones and analyzed by Western blot for expression of fusion protein. While the Jnm1-3HA and Jnm1-13Myc fusion proteins were easily detected, we were unable to detect either Yll049w-3HA or Yll049-13Myc. One possible explanation is that the expression level from the endogenous *YLL049W* promoter is so low that the fusion proteins cannot be detected. We thus obtained from Dr Heng Zhu a plasmid (with URA3 as the selectable marker) that has been reported to overexpress GST-Yll049w under control of the robust galactose-inducible *GAL1* promoter (Zhu et al. 2001).We transformed this plasmid into yeast strains expressing both Jnm1-3HA and Jnm1-13Myc and grew the transformants in synthetic medium lacking uracil (for selecting the plasmid). Standard galactose induction protocol was followed to induce expression of GST-Yll049w (Zhu et al. 2001). Again, we were unable to detect the GST-Yll049w fusion protein in these strains. In contrast, GST-Ctf4 and GST-Jnm1 fusion proteins were expressed at high levels from strains harboring the corresponding GAL1-GST fusion plasmids under the same conditions. This result suggests that the Yll039w protein might become extremely unstable when tagged with epitope tags. Given that the Yll049w fusion proteins were not expressed at

detectable level, we were unable to perform co-IP or GST pull-down experiments to confirm a physical interaction between Yll049w and Jnm1.We also note that a yeast strain expressing Yll049w-TAP was not available from the collection of TAP-tagged yeast strains made by O'Shea and Weissman's group (Ghaemmaghami et al. 2003), possibly because such a strain did not express detectable fusion protein.

Statement of contributions

PY developed statistical and computational methods and generated information-based predictions.

BDP developed the congruence calculation, conducted the nuclear migration and benomyl screens, and two-hybrid test for interaction between Yll049wp and Jnm1p.

XP conducted the *PFD1*, *LTE1*, *SPO12*, and *SLK19* dSLAM screens and the suppression of synthetic lethality between *LTE1* and the Sin3/Rpd3 components by

TAB1-6.

JDB and FAS helped initiate and supervised the experimental work.

FAS and JSB helped initiate the theoretical work, and JSB supervised the theoretical and computational work.

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25–9.

Boutros M, Kiger AA, Armknecht S, Kerr K, Hild M, Koch B, Haas SA, Consortium HF, Paro R, Perrimon N. 2004. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 303(5659):832–5.

Carvalho P, Gupta Jr ML, Hoyt MA, Pellman D. 2004. Cell cycle control of kinesin-mediated transport of Bik1 (CLIP-170) regulates microtubule stability and dynein activation. *Dev Cell* 6(6):815–29.

Farkasovsky M, Kuntzel H. 2001. Cortical Num1p interacts with the dynein intermediate chain Pac11p and cytoplasmic microtubules in budding yeast. *J Cell Biol* 152(2):251–62.

Fleming JA, Vega LR, Solomon F. 2000. Function of tubulin binding proteins *in vivo*. *Genetics* 156(1):69–80.

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick

JM, Michon AM, Cruciat CM, et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415(6868):141–7.

Geissler S, Siegers K, Schiebel E. 1998. A novel protein complex promoting formation of functional alpha- and gamma-tubulin. *EMBO J* 17(4):952–66.

Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. *Nature* 425(6959):737–41.

Gheber L, Kuo SC, Hoyt MA. 1999. Motile properties of the kinesin-related Cin8p spindle motor extracted from *Saccharomyces cerevisiae* cells. *J Biol Chem* 274(14):9564–72.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418(6896):387–91.

Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, et al. 2003. A protein interaction map of *Drosophila melanogaster*. *Science* 302(5651):1727–36.

Goehring AS, Mitchell DA, Tong AH, Keniry ME, Boone C, Sprague Jr GF. 2003. Synthetic lethal analysis implicates Ste20p, a p21-activated protein kinase, in polarisome activation. *Mol Biol Cell* 14(4):1501–16.

Goldberg DS, Roth FP. 2003. Assessing experimentally derived interactions in a small world. *Proc Natl Acad Sci USA* 100(8):4372–6.

Gunsalus KC, Yueh WC, MacMenamin P, Piano F. 2004. RNAiDB and PhenoBlast: web tools for genome-wide phenotypic mapping projects. *Nucleic Acids Res* 32(Database issue):D406–10.

Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, et al. 2002. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415(6868):180–3.

Hoyt MA, Macke JP, Roberts BT, Geiser JR. 1997. *Saccharomyces cerevisiae* PAC2 functions with CIN1, 2 and 4 in a pathway leading to normal microtubule stability. *Genetics* 146(3):849–57.

Hoyt MA, Stearns T, Botstein D. 1990. Chromosome instability mutants of *Saccharomyces cerevisiae* that are defective in microtubule-mediated processes. *Mol Cell Biol* 10(1):223–34.

Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci USA* 98(8):4569–74.

James P, Halladay J, Craig EA. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144(4):1425–36

Kahana JA, Schlenstedt G, Evanchuk DM, Geiser JR, Hoyt MA, Silver PA. 1998. The yeast dynactin complex is involved in partitioning the mitotic spindle between mother and daughter cells during anaphase B. *Mol Biol Cell* 9(7):1741–56.

Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot

137

N, Moreno S, Sohrmann M, et al. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421(6920):231–7.

Kelley R, Ideker T. 2005. Systematic interpretation of genetic interactions using protein networks. *Nat Biotechnol* 23(5):561–6.

Knaus M, Cameroni E, Pedruzzi I, Tatchell K, De Virgilio C, Peter M. 2005. The Bud14p–Glc7p complex functions as a cortical regulator of dynein in budding yeast. *EMBO J* 24(17):3000–11.

Lee WL, Oberle JR, Cooper JA. 2003. The role of the lissencephaly protein Pac1 during nuclear migration in budding yeast. *J Cell Biol* 160(3):355–64.

Li J, Lee WL, Cooper JA. 2005. NudEL targets dynein to microtubule ends through LIS1. *Nat Cell Biol* 7(7):686–90.

Loewith R, Smith JS, Meijer M, Williams TJ, Bachman N, Boeke JD, Young D. 2001. Pho23 is associated with the Rpd3 histone deacetylase and is required for its normal function in regulation of gene expression and silencing in *Saccharomyces cerevisiae*. *J Biol Chem* 276(26):24068–74.

Longtine MS, McKenzie III A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14(10):953–61.

Lord PW, Stevens RD, Brass A, Goble CA. 2003. Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. *Bioinformatics* 19(10):1275–83.

Mewes HW, Amid C, Arnold R, Frishman D, Guldener U, Mannhaupt G, Munsterkotter M, Pagel P, Strack N, Stumpflen V, et al. 2004. MIPS: analysis and annotation of proteins from whole genomes. *Nucleic Acids Res* 32(Database issue):D41–4.

Miller RK, Heller KK, Frisen L, Wallack DL, Loayza D, Gammie AE, Rose MD. 1998. The kinesin-related proteins, Kip2p and Kip3p, function differently in nuclear migration in yeast. *Mol Biol Cell* 9(8):2051–68.

Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303(5656):343–8.

Ooi SL, Shoemaker DD, Boeke JD. 2003. DNA helicase gene interaction network defined using synthetic lethality analyzed by microarray. *Nat Genet* 35(3):277–86.

Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD. 2004. A robust toolkit for functional profiling of the yeast genome. *Mol Cell* 16(3):487–496

Schlitt T, Palin K, Rung J, Dietmann S, Lappe M, Ukkonen E, Brazma A. 2003. From gene networks to gene function. *Genome Res* 13(12):2568–76.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–504.

Sheeman B, Carvalho P, Sagot I, Geiser J, Kho D, Hoyt MA, Pellman D. 2003.

Determinants of *S. cerevisiae* dynein localization and activation: implications for the mechanism of spindle positioning. *Curr Biol* 13(5):364–72.

Shou W, Sakamoto KM, Keener J, Morimoto KW, Traverso EE, Azzam R, Hoppe GJ, Feldman RM, DeModena J, Moazed D, et al. 2001. Net1 stimulates RNA polymerase I transcription and regulates nucleolar structure independently of controlling mitotic exit. *Mol Cell* 8(1):45–55.

Stegmeier F, Visintin R, Amon A. 2002. Separase, polo kinase, the kinetochore protein Slk19, and Spo12 function in a network that controls Cdc14 localization during early anaphase. *Cell* 108(2):207–20.

Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, et al. 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294(5550):2364–8.

Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, et al. 2004. Global mapping of the yeast genetic interaction network. *Science* 303(5659):808–13.

Tucker CL, Fields S. 2003. Lethal combinations. *Nat Genet* 35(3):204–5.

Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, et al. 2000. A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* 403(6770):623–7.

Vainberg IE, Lewis SA, Rommelaere H, Ampe C, Vandekerckhove J, Klein HL, Cowan NJ. 1998. Prefoldin, a chaperone that delivers unfolded proteins to cytosolic

chaperonin. *Cell* 93(5):863–73.

van Haaften G, Vastenhouw NL, Nollen EA, Plasterk RH, Tijsterman M. 2004.
Gene interactions in the DNA damage-response pathway identified by genome-wide
RNA-interference analysis of synthetic lethality. *Proc Natl Acad Sci USA*
101(35):12992–6.

Wong SL, Zhang LV, Tong AH, Li Z, Goldberg DS, King OD, Lesage G, Vidal M,
Andrews B, Bussey H, et al. 2004. Combining biological networks to predict genetic
interactions. *Proc Natl Acad Sci USA* 101(44):15682–7.

Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R,
Bidlingmaier S, Houfek T, et al. 2001. Global analysis of protein activities using
proteome chips. *Science* 293(5537):2101–5.

Figure 1. Congruent synthetic lethal (SL) interactions are consistent with functional pathway membership. (A) A simplified synthetic lethality pathway model. Black arrows indicate the schematic flow of a process, with essential genes (red circles) connected by non-essential genes (black circles) organized into two parallel pathway branches (black dashed lines). If at least one of the pathway branches is required for viability, SL interactions (red lines) will be observed between the pathway branches but not within a pathway branch. In this picture, deleting any component of a pathway branch destroys its activity. (B) Directly observed SL genetic interactions bridge pathway branches. The table indicates that SL interactions will be observed between components of the two pathway branches, whereas no interactions will be observed within a branch. (C) Functional associations inferred from the congruence score (blue lines) join the components of a pathway branch. The table indicates raw number of SL interaction partners shared by a pair of genes and its conversion to the congruence score, calculated as the $-\log_{10}$ P-value for partner sharing. The congruence connections are orthogonal to the direct SL interactions and align with pathway membership.

Figure 1.

A
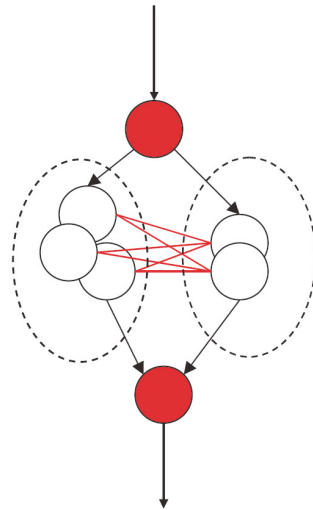
Same functional pathway:
physical interactions enhanced
SL interactions reduced
SL interaction pattern enhanced

● = essential gene
○ = non-essential gene
— = synthetic lethality (SL)
↓ = information flow
◌ = functional pathway
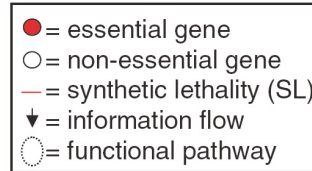
B Synthetic lethality links:
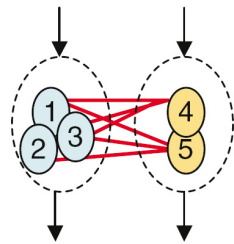
Targets

	1	2	3	4	5
1				SL	SL
2				SL	SL
3				SL	SL
4	SL	SL	SL		
5	SL	SL	SL		

Queries

C Congruence links:
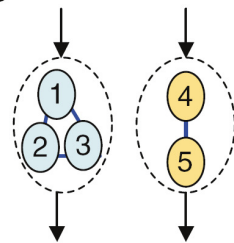
Target pairs	Shared SL interactions	Congruence score ($-\log_{10} P$)
1–2	2	4
1–3	2	4
1–4	0	0
1–5	0	0
2–3	2	4
2–4	0	0
2–5	0	0
3–4	0	0
3–5	0	0
4–5	3	6

143

Figure 2. Genetic congruence predicts physical colocalization and shared gene function. Cumulative bins were constructed for all target gene pairs using a threshold congruence score. (A) High congruence score predicts protein complex membership. The red dot at congruence score 5 indicates the threshold at which congruent gene products are more likely than synthetic lethal partners to reside in the same protein complex ($P < 10^{-6}$ for co-residence of the congruent pairs). (B) A high congruence score predicts GO annotation correlations (biological process, molecular function, and cellular component). Red symbols label the thresholds above which annotations of congruent gene pairs are more highly correlated than annotations for synthetic lethal pairs. (C) High congruence excludes synthetic lethal interaction. The black dashed line labels the threshold value of congruence score 14, above which the binomial P-value for observed number of synthetic lethal interactions is insignificant ($P > 0.05$). (D) Synthetic lethal interactions have been used to calculate congruence scores (blue lines, threshold congruence score ≥ 10) that connect genes in the same pathway branch. Congruence edges are generally orthogonal to the underlying synthetic lethal interactions and parallel to protein complex membership (green lines, membership in a single complex; black lines, overlap of congruence and protein complex edge). The shaded inset shows the synthetic lethal interactions (red lines) underlying the congruence edge between *UBA4* and *ELP6*. Congruence networks at thresholds 8 and 15 are shown as Supplementary Figures S7 and S8.

Figure 2.

145
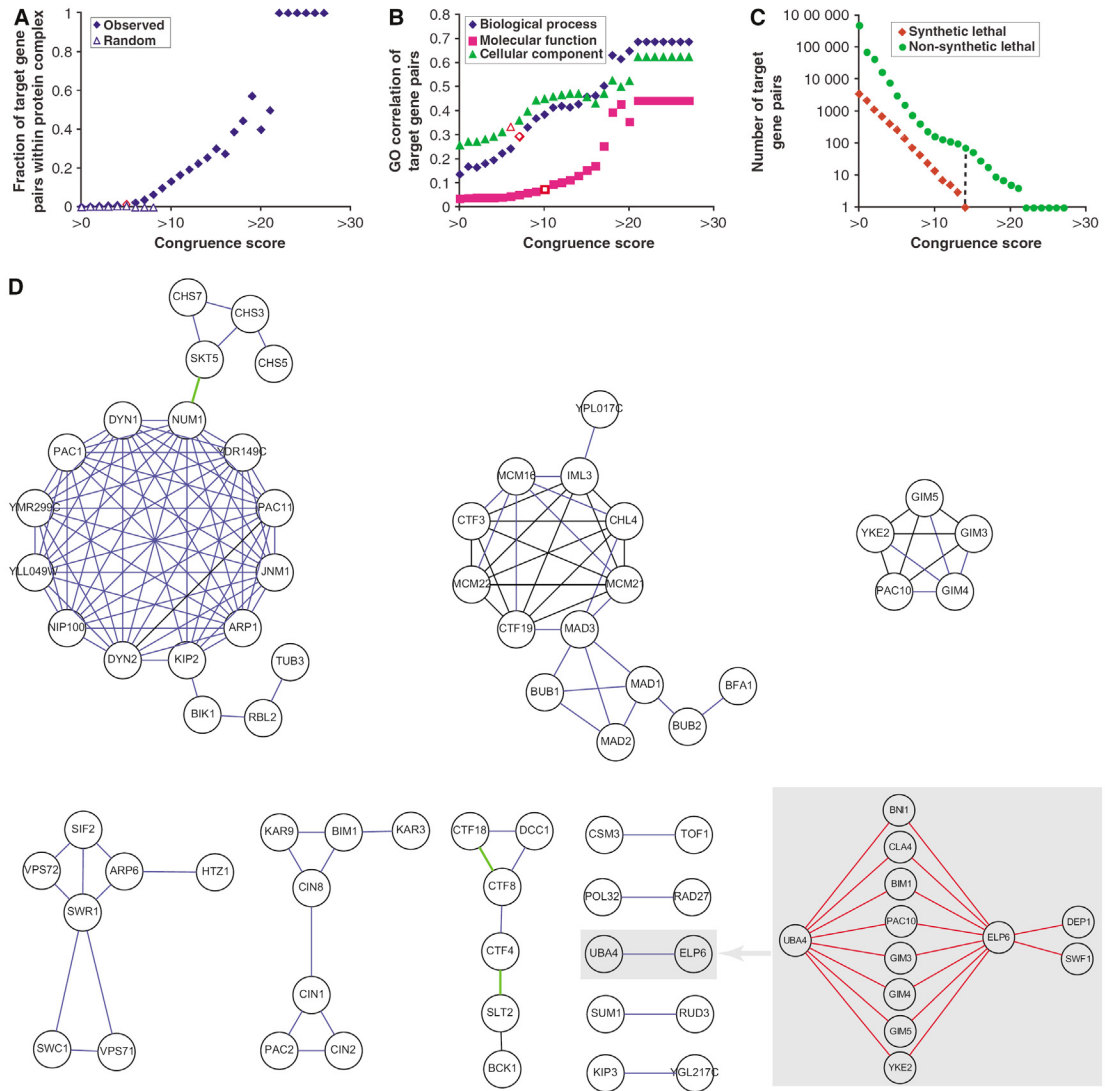
Figure 3. The congruence score but not the number of synthetic lethal interactions predicts numeric phenotypes for deletion mutants. (A) Null mutants of 59 genes with congruence score ≥ 4 for six landmark genes (*DYN1*, *ARP1*, *DYN2*, *JNM1*, *NUM1*, and *NIP100*) known to be required for robust nuclear migration were measured for percent abnormal nuclear migration at 13 ºC. Each mutant's nuclear migration defect is plotted by congruence score to each landmark gene (congruence score range is labeled) and by average congruence score (dots). (B) Null mutants of 31 genes with congruence score ≥ 4 for landmark gene *CIN1* known to be required for benomyl resistance were tested for benomyl sensitivity at concentration 5 µg/ml. The fraction of benomyl-sensitive null mutants is plotted with each congruence score cutoff. (C, D) Null mutants of 451 candidate benomyl-resistant genes are ranked based on their average congruence score or number of synthetic lethal interactions with seven landmark genes (*CIN1*, *YML094C-A*, *PAC10*, *PFD1*, *GIM3*, *TUB3*, and *GIM5*) known to be required for benomyl resistance (Pan et al., 2004). The LD_{50} benomyl concentration is defined by the lowest benomyl concentration when the control/experimental hybridization signal concentration ≥ 2. The red mark represents the median LD_{50} benomyl concentration.

Figure 3.
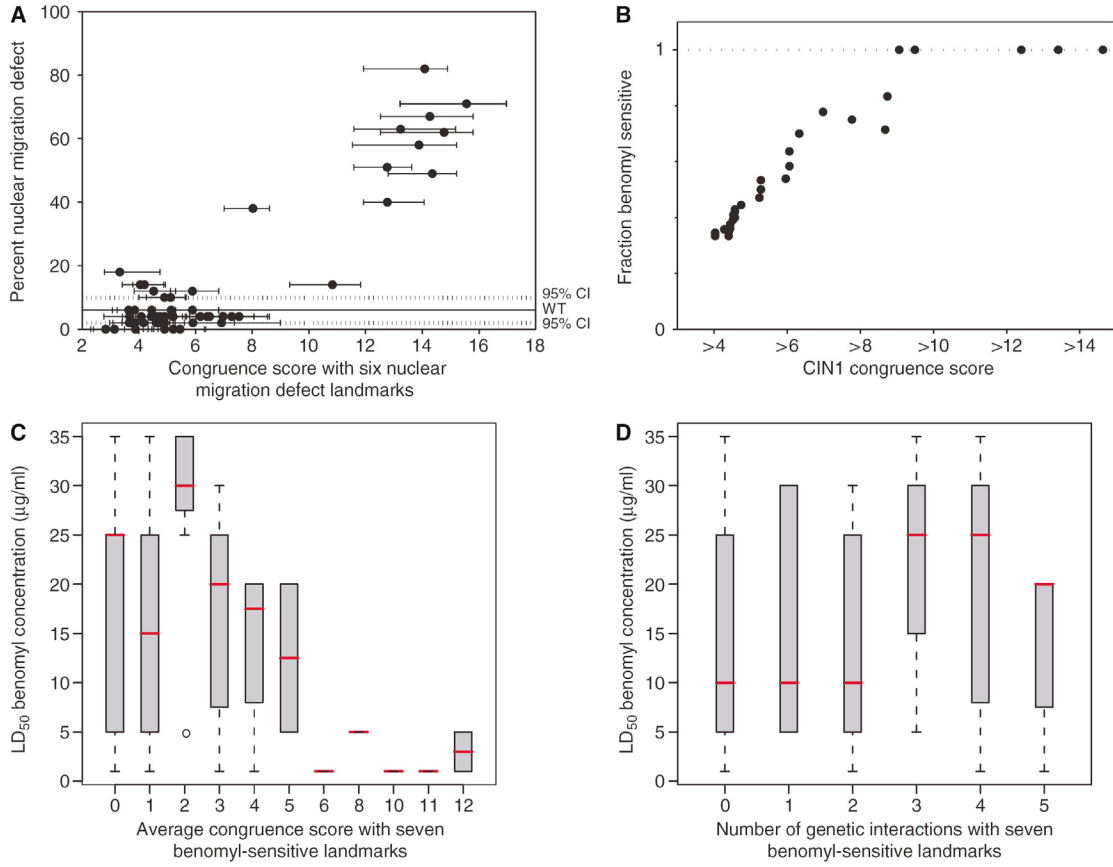
Figure 4. Parallel pathways required for mitotic exit. (A) Synthetic lethal

interactions define parallel pathways: FEAR, MEN, and Sin3/Rpd3 complex. The

Sin3/Rpd3 complex could function in the FEAR pathway; here, we depict a separate

pathway because we demonstrate that it is parallel to MEN. (B) Suppression of *rpd3Δ*

lte1Δ synthetic lethality by *TAB1-6* (*CDC14* allele). A haploid convertible heterozygous

diploid double mutant (*LTE1/lte1Δ::natMX XXX/xxxΔ::kanMX*, *XXX* stands for *RPD3* or

SLK19) was transformed with a vector (YCplac33), or a plasmid expressing the wild-type

CDC14 or *TAB1-6*. The resultant transformants were sporulated and the meiotic

progenies were spotted with 10× serial dilutions onto haploid selection media specific for

haploid *MAT*a cells of indicated genotypes. Cells were incubated at 30 ºC for 3 days and

photographed. The synthetic lethality between *LTE1* and other components of the

Sin3/Rpd3 complex was also similarly confirmed. The results of this experiment were

recapitulated by tetrad analysis on YPD (data not shown).

Figure 4.

A

FEAR   MEN   Sin3/Rpd3 complex

SPO12 ─ LTE1 ─ SAP30
SLK19 ─ CLA4 ─ RPD3 SIN3 ─ PHO23

CDC14 Release of Cdc14 from nucleolus

● = essential gene
○ = non-essential gene
— = synthetic lethality (SL)
↓ = information flow
◌ = functional pathway

Mitotic exit

B

	MATa xxxΔ::kanMX URA3 plasmid	MATa xxxΔ::kanMX lte1Δ::natMX URA3 plasmid	MATa xxxΔ::kanMX lte1Δ::natMX
rpd3Δ Vector			
rpd3Δ CDC14			
rpd3Δ TAB1-6			
slk19Δ Vector			
slk19Δ CDC14			
slk19Δ TAB1-6			
	MM–Ura +G418	MM–Ura + G418 + Nat	MM + 5-FOA + G418 + Nat

149

Supplementary information:

Synthetic lethal genes bridge parallel pathways

Protein complexes are often the functional units that implement biological processes. Knowledge of protein complex organization can help explain the functions of genes within the context of biological pathways. We hypothesized that protein complex data can reveal quantitative, hierarchical organization of synthetic lethal interactions. Specifically, members of different protein complexes in parallel pathways should cluster in groups of direct synthetic lethal partners. Synthetic lethal interactions between these groups should 'bridge' the parallel pathway branches they reveal.

To explore this hypothesis, we first calculated the total number of synthetic lethal interactions between protein complexes using synthetic lethal dataset generated from the SGA approach (Tong et al. 2004) and high-throughput protein complex datasets (Gavin et al. 2002; Ho et al. 2002). Of 3799 synthetic lethal pairs of knock-out mutants, 1083 (~30%) bridge distinct protein complexes, with one member of a synthetic lethal pair in a different complex than its partner. Since only ~1% of synthetic lethal pairs reside within the same protein complex (Tong et al., 2004), a synthetic lethal interaction is 30× more likely to bridge two distinct complexes than reside within a single complex. Analysis using interactions from curated protein complex data (Mewes et al. 2004) and from high-throughput yeast two-hybrid screens (Ito et al. 2001; Uetz et al. 2000) also support our hypothesis that synthetic lethal pairs are more likely to encode proteins without direct physical interactions. A recent computational study reports similar results (Wong et al. 2004).

For each pair of protein complexes reported in large-scale screens (Gavin et al. 2002; Ho et al. 2002), enrichment of synthetic lethal interaction was quantified as the parallel complex score calculated as the negative $\log_{10}$ of the binomial *p*-value for number of synthetic lethal interactions observed between members of the two complexes given the overall frequency of synthetic lethal interactions observed in the whole data set (see the next section for detail). Significant numbers of protein complex pairs are observed being bridged by synthetic lethal interactions using the actual synthetic lethal interactions compared to a randomized set when parallel complex scores ≥ 3 (*p*-value $< 10^{-5}$, Fig. S1A). The hierarchical view of synthetic lethal interaction by clustering gene products into protein complexes shows protein complex nodes connected by highly significant parallel complex linkages (Fig. S1B). Analysis using the curated MIPS protein complex dataset (Mewes et al. 2004) generates similar results (Fig. S2).

The PAC10 complex is the hub of the parallel complex network, with links to 34 other protein complexes (Fig. S1B). Its hub character is due in part to the bias that all four complex components, *PAC10*, *GIM3*, *GIM5*, and *YKE2* (Gavin et al., 2002), are SGA query genes (Tong et al. 2004). The PAC10 complex proteins detected by mass spectrometry belong to the biochemically characterized Prefoldin complex (*PAC10*, *GIM3*, *GIM4*, *GIM5*, *YKE2*, *PFD1*), involved in tubulin folding and delivering unfolded proteins to cytosolic chaperonin (Geissler et al. 1998; Vainberg et al. 1998). Deletion mutants of Prefoldin complex components are viable, and sensitive to the microtubule depolymerizing drug benomyl (Geissler et al. 1998). The 34 protein complexes linked with the PAC10 complex carry out diverse biological processes including cytoskeleton

organization and biogenesis, budding, transcription regulation, translational membrane

targeting, rRNA processing, and DNA damage response (Ashburner et al. 2000). The

synthetic lethal interaction linkages indicate that these pairs of protein complexes provide

related, but distinct, cellular functions. For some linkages, the relationship is readily

understood given current knowledge. For example, the PAC10 complex exhibits

enhanced synthetic lethal interactions with the IML3 complex (*IML3*, *MCM21*, *MCM22*,

CTF3, *CTF19*, *CHL4*, *AME1*, *NKP1*) (Fig. S1B), which is a kinetochore component. It is

reasonable to propose that activities of these two protein complexes may be

complementary during kinetochore capture or during chromosome movement, when

microtubule dynamics and kinetochore activity are coupled. Synthetic lethality may be

explained by higher-order effects of combined perturbations of microtubules and

kinetochores.

Probability of synthetic lethal interaction and parallel complex score

The probability of at least k synthetic lethal interactions bridging two protein

complexes was calculated from the binomial distribution:

$$p(x \geq k) = \sum_{x=k}^{n} C(n,x) \cdot P^{x} \cdot (1-P)^{n-x} \quad ,$$

in which $C(n,x)$ is the combinatorial factor $n!/x!(n-x)!$; n is the total number of possible

interactions between two protein complexes; and k is the number of observed synthetic

lethal interactions between two protein complexes. The probability of observing a set of

synthetic lethal interactions between two protein complexes P was approximated to be

0.0064 from a/bc, where b equals 126, the number of query genes, c equals 4700, the

number of target genes, and *a* equals 3799, the number of total synthetic lethal interactions observed between query and target genes. The parallel complex score is $-\log_{10}[p(x \geq k_{obs})]$.

We reasoned that for a final *p*-value of 0.01, an appropriate single-test *p*-value that incorporates multiple testing all pairs of 780 protein complexes used would be $\sim 0.01/7802 = 2 \times 10^{-8}$, corresponding to a parallel complex score of 7 to 8. As the number of complex pairs nearly doubles when the parallel complex score decreases from 8 to 7 (Fig. S1A), we used 8 as the threshold for the visualization (Fig. S1B). Similarly, threshold 7 was used for Fig. S2B.

Protein complex pair sharing protein components

The Jaccard coefficient $c = (n1 \cap n2)/(n1 \cup n2)$, where *n1* is the number of proteins in complex 1 and *n2* is the number of proteins in complex 2, was calculated to define comparable protein complexes. The value of 0.4 was used as the threshold of Jaccard coefficient to define similar complexes in Fig. S1B.

Figure S1. Synthetic lethal genes bridge parallel pathways from analysis on high throughput protein complex dataset (Gavin et al. 2002; Ho et al. 2002). (A) Significant numbers of protein complexes are bridged by synthetic lethal interactions than expected by chance (p-value $< 10^{-5}$ when parallel complex score $\geq$ 3). (B) Pairwise synthetic lethal interactions have been mapped to the level of protein complexes (circles) using the parallel complex score with threshold value $\geq$ 8 (red lines). The size of a circle indicates the number of proteins in the complex, and its color indicates the number of corresponding genes used as SGA queries. Independently reported protein complexes that share multiple components (Jaccard coefficient $\geq$ 0.4) are linked (dashed black lines). The shaded inset depicts the pairwise synthetic lethal interactions between components of the PAC10 and IML3 complexes that are summarized by a single parallel complex edge.

Figure S1.

A

B

155
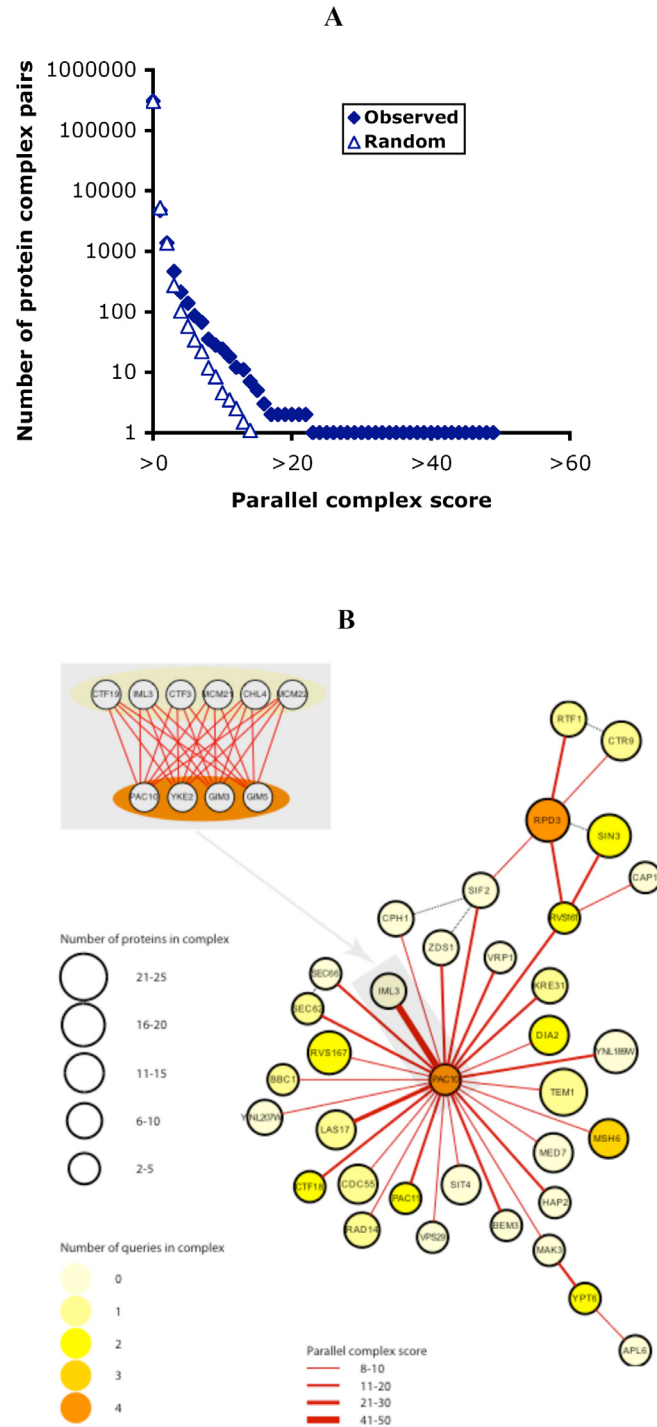
Figure S2. Synthetic lethal genes bridge parallel pathways from analysis on curated MIPS protein complex dataset (Mewes et al. 2004). (A) Significant numbers of protein complexes are bridged by synthetic lethal interactions than expected by chance (p-value < 0.001 when parallel complex score $\geq$ 5). (B) Pairwise synthetic lethal interactions have been mapped to the level of protein complexes (circles) using the parallel complex score with threshold value $\geq$ 7 (red lines).

Figure. S2.

A

B
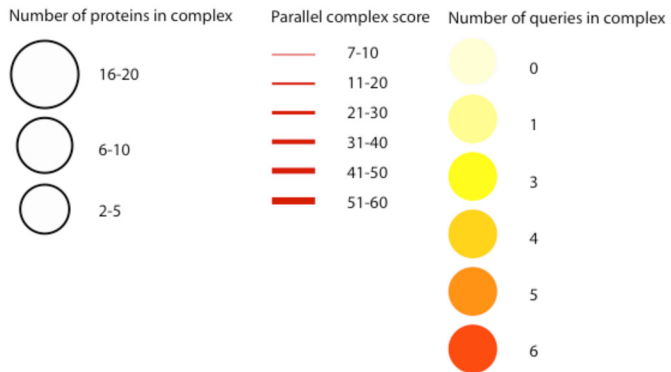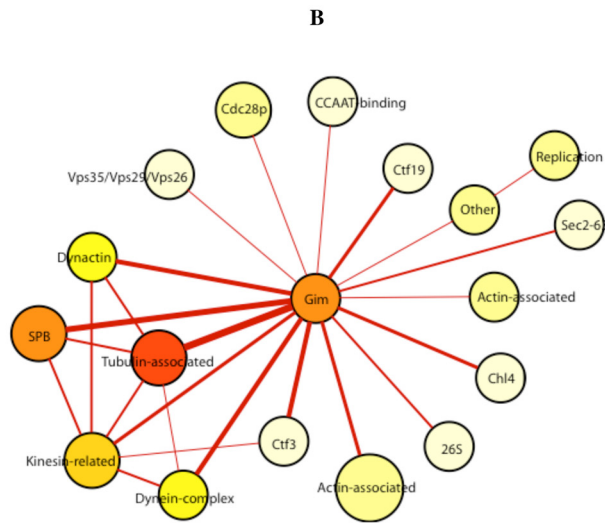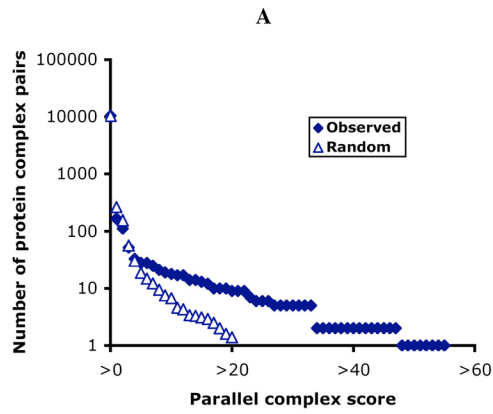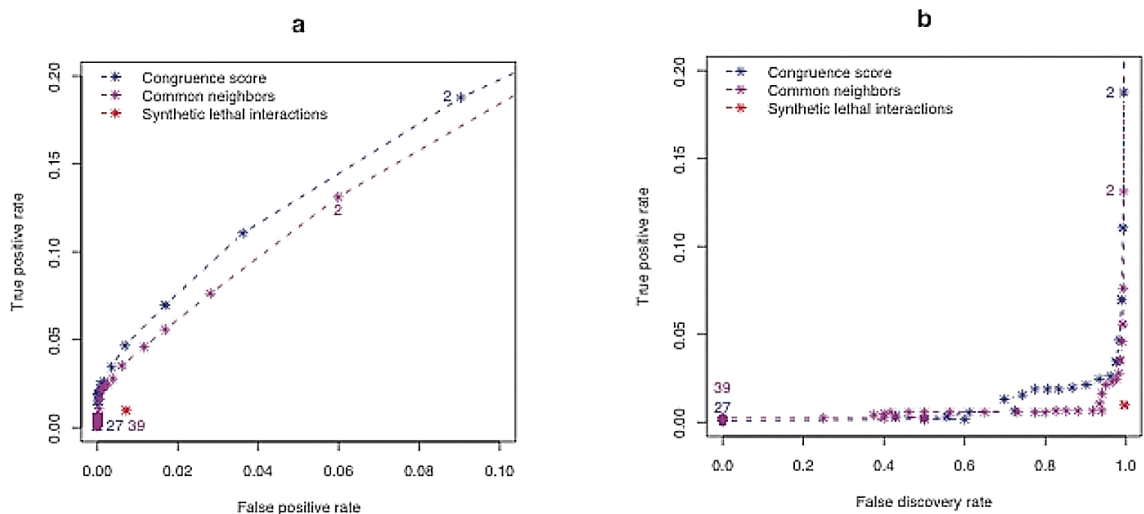
Figure S3. The congruence score method is superior to the number of common neighbors in predicting protein complex coresidence of congruent gene encoded proteins. Prediction of coresidence is presented as a receiver operating characteristic (ROC) curve in terms of the false-positive rate (A), equal to (false positives) / (false positives + true negatives), and the false-discovery rate (B), equal to (false positives) / (false positives + true positives). The numbers indicate the cut-off values for congruence score (blue) and common neighbors (purple). Synthetic lethal interaction (red) has higher false positive rate (A) and higher false discovery rate (B) in predicting protein complex coresidence as compared with congruence score method when their true positive rates are comparable. The higher ordinate for the congruence score method implies superior performance based on the area under the curve criterion.

158
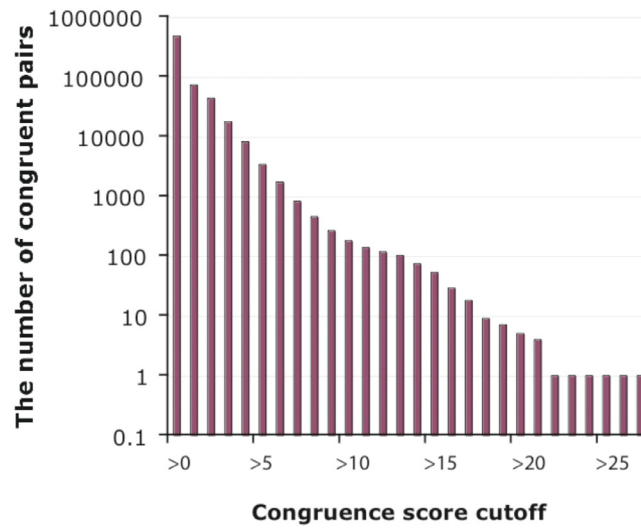
Figure S4. The number of target congruent gene pairs at each congruence score cut-off.
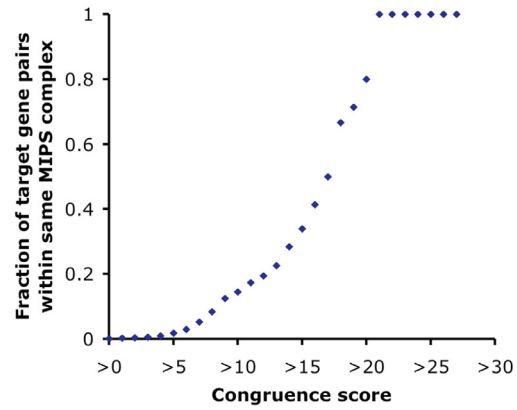
Figure S5. Congruence score predicts protein complex membership using curated MIPS protein complex dataset.
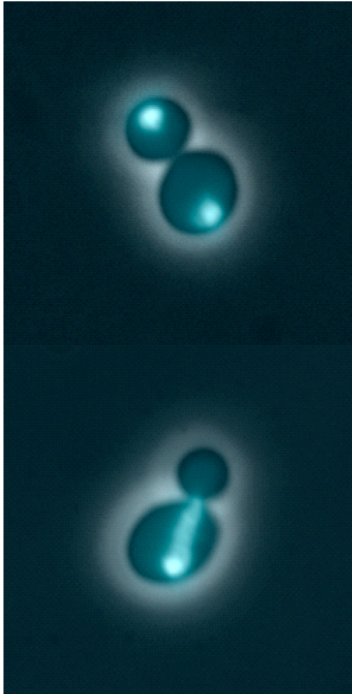
Normal: Abnormal:
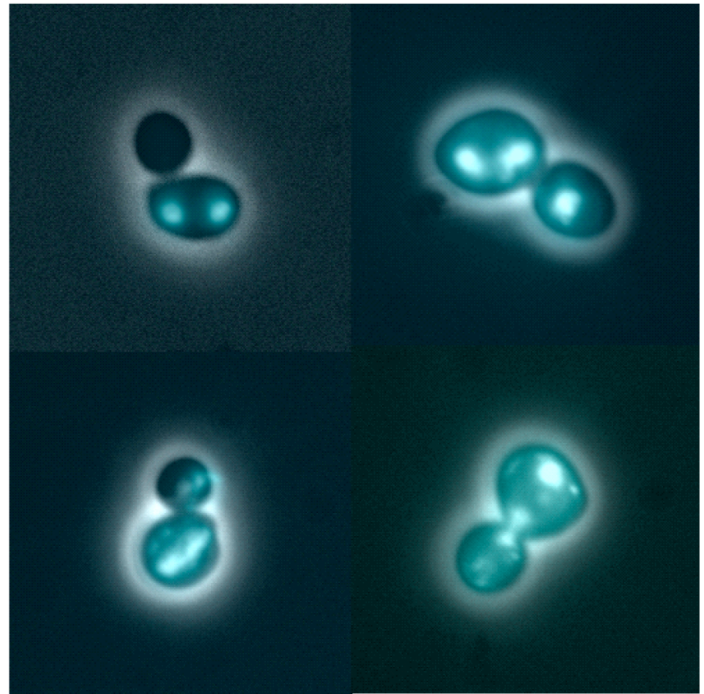
Figure S6. Examples of nuclear migration phenotypes. Left panel, merged

Phase/DAPI images of normal nuclear migration events; right panel, merged Phase/DAPI

images of abnormal nuclear migration events.

Figure S7. Jnm1p binds Yll049wp by yeast-two hybrid assay.

Yeast-two-hybrid experiments were performed using activation and binding domain vectors pOAD (*LEU2*-marked) and pOBD-2 (*TRP1*-marked), respectively, and yeast strains PJ69-4a and PJ69-4α (James et al. 1996). Materials were kindly provided by Stanley Fields, Yeast Resource Center. GAL4-binding domain fusions were transformed into PJ69-4α and GAL4-activation domain fusions were transformed into PJ69-4a. The two strains were mated and diploids were selected on SC −Leu −Trp. The resulting diploids were plated on SC −Ade −His media in two dilutions (2 µl of 0.1 OD_{600}/ml and 0.02 OD_{600}/ml) at 30 °C. Growth at 4 days is shown.

JNM1-BD, *JNM1* fusion with GAL4 binding domain; YLL049W-AD, *YLL049W* fusion with GAL4 activation domain; BD, binding domain alone; AD, activation domain alone.

Two independent JNM1-BD and two independent YLL049W-AD transformants supported growth when appropriately combined. YLL049W-BD + AD alone resulted in growth due to self-activation and was therefore not informative (data not shown).

The plasmids and strains used for this study are distinct from those used by Ito, et al. (Ito et al. 2001), who reported high-throughput yeast-two-hybrid interaction between *JNM1* and *YLL049W*.

Genotypes: PJ69-4a: *MAT*a *trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ* PJ69-4alpha: *MAT*α *trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*
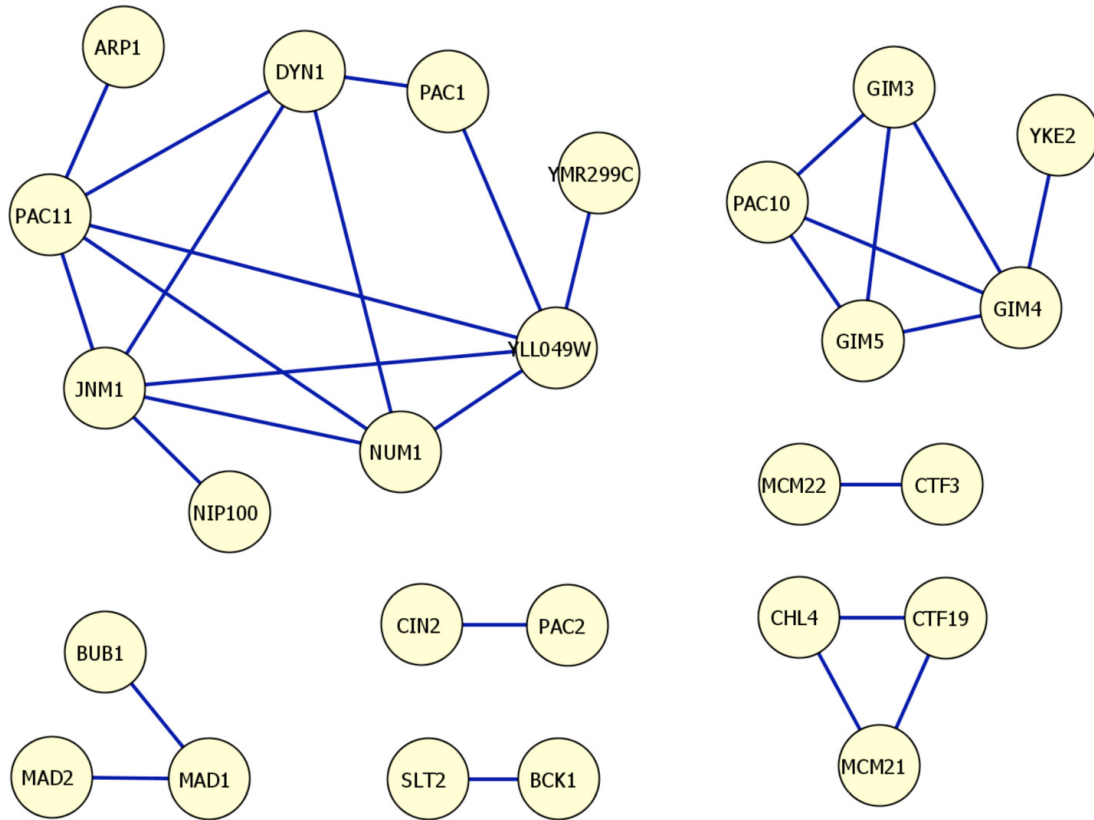
Figure S7.
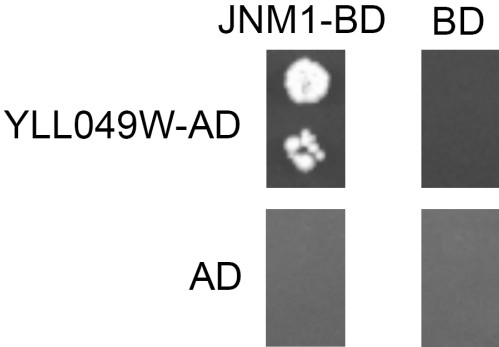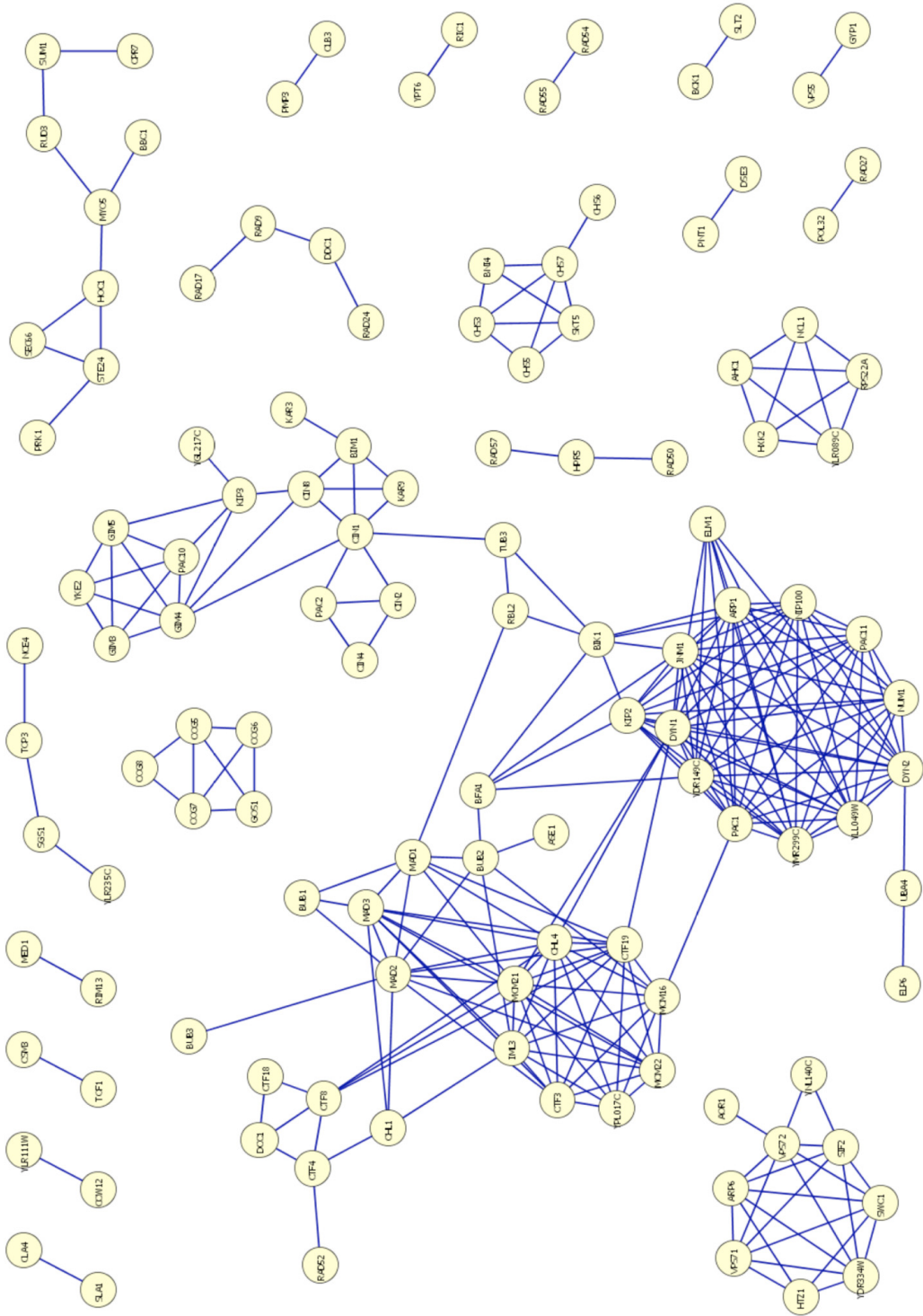
Figure S8. Target gene pair congruence network with the congruence score cutoff greater than or equal to 8.
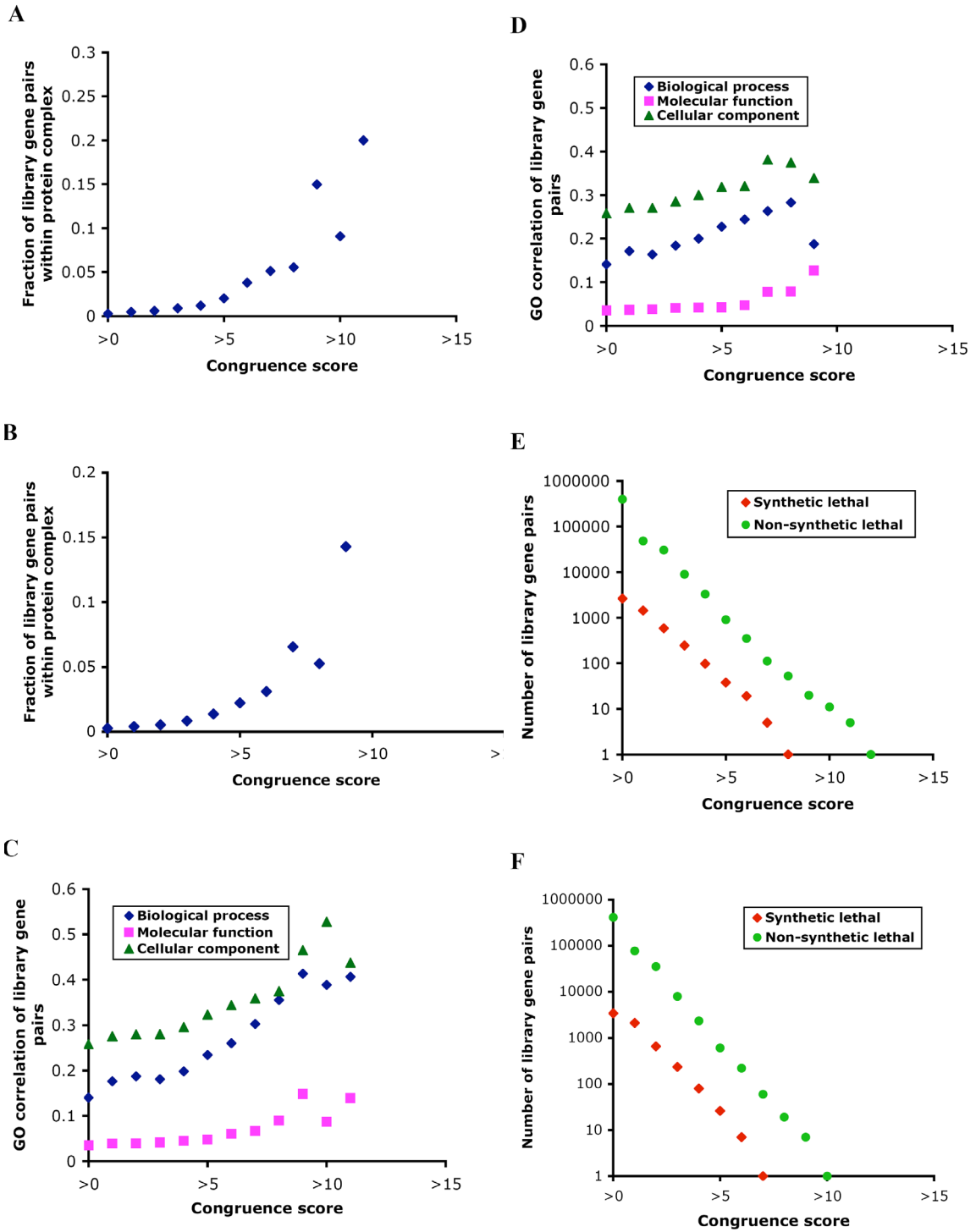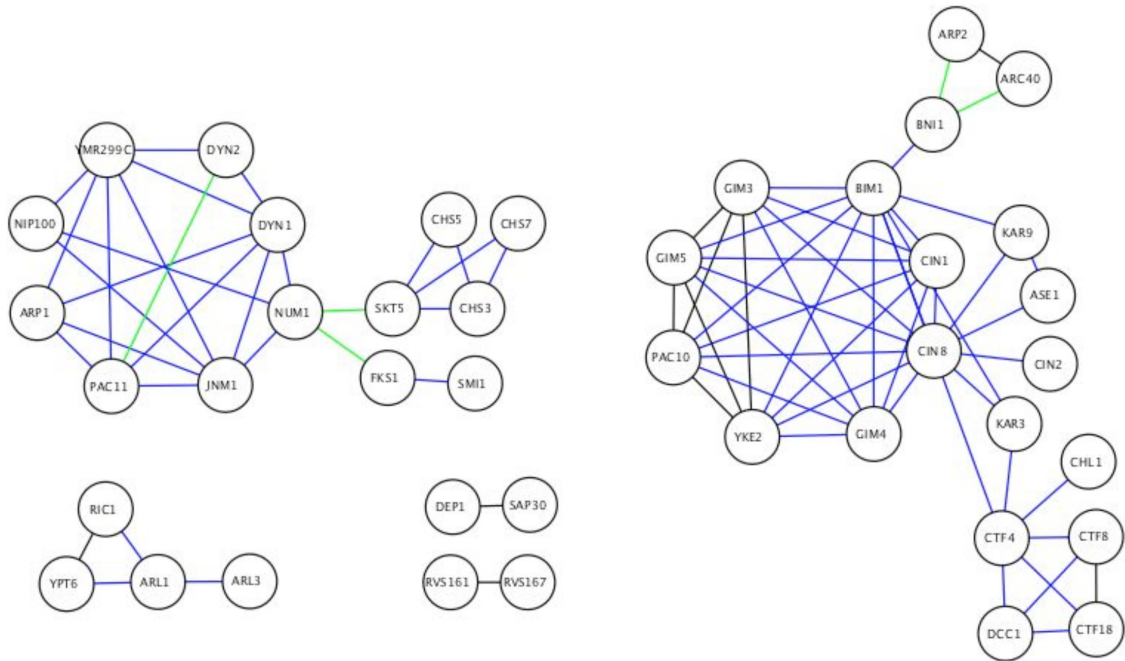
Figure S9. Target gene pair congruence network with the congruence score cutoff greater than or equal to 15.

Figure S10. Noise robustness analysis for the congruence method. (A), (C), and (E) are results derived from the dataset containing 30% of false negative synthetic lethal interactions. (B), (D), and (F) are results derived from the dataset containing 30% of false positive synthetic lethal interactions. The congruence scores generated from datasets containing false negatives and false positives are in the range of 0 to 10 and show similar results as those using the original dataset (compare with Fig. 2a, 2b, 2c). (A) and (B) A high congruence score predicts protein complex membership. Above congruence score of 3, significant numbers of congruence gene products reside in the same complex as compared with synthetic lethal gene products ($P < 0.05$). (C) and (D) A high congruence score predicts Gene Ontology (GO) annotation correlations (biological process, molecular function, and cellular component). Above congruence score of (7, 6, 5) and (8, 7, 5) for (c) and (d), respectively, congruence pairs have significantly higher GO correlation (biological process, molecular function, cellular component) as compared with that of synthetic lethal gene pairs ($P < 0.05$). (E) and (F) High congruence excludes synthetic lethal interaction. Above congruence score of 8 and 7 for (E) and (F), respectively, the binomial p-value for observed number of synthetic lethal interactions is insignificant ($P > 0.05$).

166

Figure S10.

Figure S11. Query gene pair genetic congruence network with the congruence score cutoff greater than or equal to 33. Congruent interactions are labeled with blue lines, physical interactions derived from any two proteins in the same protein complex are labeled with green lines, and black lines represent coexistent congruent and physical interactions.
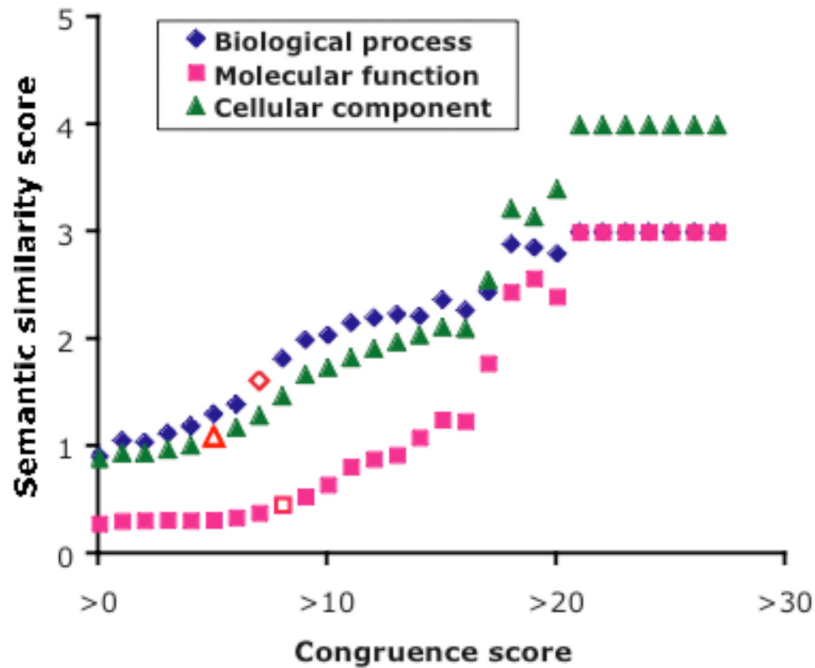
Figure S12. Semantic similarity of congruent genes.

Semantic similarity (Lord et al. 2003) was calculated for congruent gene pairs and synthetic lethal gene pairs using all yeast gene Gene Ontology annotations for training. Open points indicated the congruence scores at which the semantic similarity for congruent genes rises above similarity for synthetic lethal genes (significance $p < 0.05$). These points show that congruence score out-perform direct synthetic lethal interactions at thresholds of 7 (process), 8 (function), and 5 (component). This is superior performance to that indicated in the main text using GO depth correlation, where the crossovers occurred at 7 (process), 10 (function), and 6 (component).

169

Table S1. True positive rate and false positive rate using different threshold values for congruence score method and number of common neighbors in predicting protein complex coresidence.

	Congruence method				Number of common neighbors			
	TP number	FP number	TP rate	FP rate	TP number	FP number	TP rate	FP rate
0	1220	480451	1	1	1220	480451	1	1
1	307	72444	0.2516	0.1508	307	72611	0.2516	0.1511
2	229	43433	0.1877	0.0904	160	28751	0.1311	0.0598
3	135	17388	0.1107	0.0362	93	13502	0.0762	0.0281
4	85	8102	0.0697	0.0169	68	8104	0.0557	0.0169
5	57	3318	0.0467	0.0069	56	5578	0.0459	0.0116
6	42	1687	0.0344	0.0035	43	2924	0.0352	0.0061
7	32	790	0.0262	0.0016	34	1846	0.0279	0.0038
8	30	423	0.0246	0.0009	30	1236	0.0246	0.0026
9	26	238	0.0213	0.0005	28	793	0.023	0.0017
10	24	156	0.0197	0.0003	26	493	0.0213	0.001
11	23	115	0.0189	0.0002	20	315	0.0164	0.0007
12	23	95	0.0189	0.0002	13	186	0.0107	0.0004
13	23	79	0.0189	0.0002	8	125	0.0066	0.0003
14	19	55	0.0156	0.0001	8	89	0.0066	0.0002
15	16	37	0.0131	0.0001	8	59	0.0066	0.0001
16	8	21	0.0066	0	8	46	0.0066	0.0001
17	7	11	0.0057	0	8	38	0.0066	0.0001
18	4	5	0.0033	0	7	28	0.0057	0.0001
19	4	3	0.0033	0	7	24	0.0057	0
20	2	3	0.0016	0	7	19	0.0057	0
21	2	2	0.0016	0	7	13	0.0057	0
22	1	0	0.0008	0	7	9	0.0057	0
23	1	0	0.0008	0	7	7	0.0057	0
24	1	0	0.0008	0	7	6	0.0057	0
25	1	0	0.0008	0	7	5	0.0057	0
26	1	0	0.0008	0	6	4	0.0049	0
27	1	0	0.0008	0	5	3	0.0041	0
28	-	-	-	-	4	3	0.0033	0
29	-	-	-	-	3	3	0.0025	0
30	-	-	-	-	3	3	0.0025	0
31	-	-	-	-	3	2	0.0025	0
32	-	-	-	-	3	1	0.0025	0
33	-	-	-	-	3	0	0.0025	0

	Congruence method				Number of common neighbors			
	TP number	FP number	TP rate	FP rate	TP number	FP number	TP rate	FP rate
34	-	-	-	-	3	0	0.0025	0
35	-	-	-	-	3	0	0.0025	0
36	-	-	-	-	1	0	0.0008	0
37	-	-	-	-	1	0	0.0008	0
38	-	-	-	-	1	0	0.0008	0
39	-	-	-	-	1	0	0.0008	0
ROC area	0.555				0.553			
SE	0.00859				0.00858			

Table S2. Nuclear migration phenotypes at 13 ºC. Deletion mutants for each gene were obtained from the yeast deletion collection (Research Genetics).

ORF name	Gene name	Normal	Abnormal	Percent abnormal	Average congruence score	GO slim Biological Process
(BY4741)	(WT)	94	6	6	NA	
YDR488C	PAC11	29	71	71	15.6	cytoskeleton organization and biogenesis
YMR294W	JNM1	38	62	62	14.8	cell cycle
YKR054C	DYN1	51	49	49	14.4	cytoskeleton organization and biogenesis
YDR150W	NUM1	33	67	67	14.3	cytoskeleton organization and biogenesis
YHR129C	ARP1	18	82	82	14.1	cell cycle
YLL049W		42	58	58	13.9	unknown
YOR269W	PAC1	37	63	63	13.2	cytoskeleton organization and biogenesis
YDR424C	DYN2	60	40	40	12.8	cytoskeleton organization and biogenesis
YMR299C		49	51	51	12.8	unknown
YPL155C	KIP2	86	14	14	10.8	cytoskeleton organization and biogenesis
YCL029C	BIK1	31	19	38	8.0	cell cycle
YKL048C	ELM1	48	2	4	7.5	cytokinesis
YJR053W	BFA1	48	2	4	7.3	conjugation
YPL018W	CTF19	48	2	4	7.0	cell cycle
YDR254W	CHL4	48	2	4	7.0	cell cycle
YDR318W	MCM21	48	2	4	7.0	cell cycle
YHR111W	UBA4	49	1	2	6.9	protein modification
YPR046W	MCM16	48	2	4	6.5	cell cycle
YBR107C	IML3	48	2	4	6.4	cell cycle
YMR055C	BUB2	48	2	4	6.2	cell cycle
YMR312W	ELP6	49	1	2	5.9	transcription
YJR135C	MCM22	47	3	6	5.9	cell cycle
YLR381W	CTF3	44	6	12	5.9	cell cycle
YOR058C	ASE1	50	0	0	5.5	cell cycle
YJL030W	MAD2	50	0	0	5.2	cell cycle
YOR265W	RBL2	48	2	4	5.2	protein binding
YLR386W	VAC14	47	3	6	5.1	organelle organization and biogenesis
YCR086W	CSM1	45	5	10	5.1	DNA metabolism
YLR089C		50	0	0	4.9	unknown
YOR023C	AHC1	50	0	0	4.9	DNA metabolism
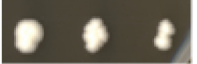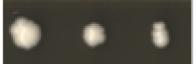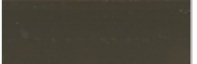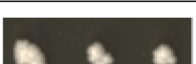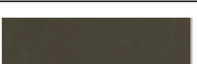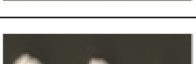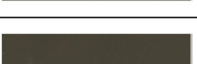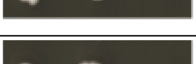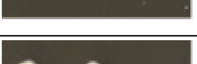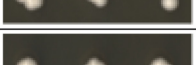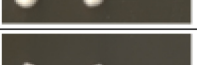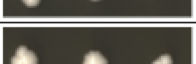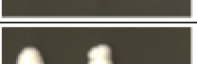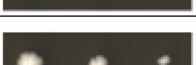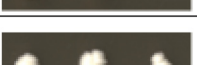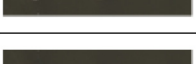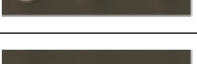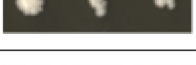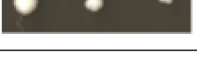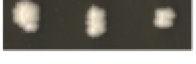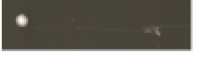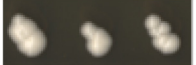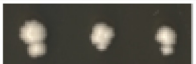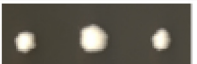
ORF name	Gene name	Normal	Abnormal	Percent abnormal	Average congruence score	GO slim Biological Process
YBL024W	NCL1	48	2	4	4.9	RNA metabolism
YBL051C	PIN4	48	2	4	4.9	cell cycle
YJL190C	RPS22A	49	1	2	4.9	protein biosynthesis
YER177W	BMH1	45	5	10	4.9	pseudohyphal growth
YML124C	TUB3	49	1	2	4.8	meiosis
YOR264W	DSE3	48	2	4	4.7	unknown
YOR266W	PNT1	48	2	4	4.7	membrane organization and biogenesis
YGL086W	MAD1	49	1	2	4.6	transport
YPL017C		48	2	4	4.6	unknown
YMR078C	CTF18	44	6	12	4.5	cell cycle
YPL008W	CHL1	47	3	6	4.4	cell cycle
YPR023C	EAF3	48	2	4	4.4	protein modification
YMR138W	CIN4	43	7	14	4.2	cytoskeleton organization and biogenesis
YIL095W	PRK1	49	1	2	4.2	cytokinesis
YCL016C	DCC1	48	2	4	4.1	cell cycle
YLR085C	ARP6	43	7	14	4.0	transport
YGR270W	YTA7	50	0	0	3.9	protein catabolism
YLL006W	MMM1	49	1	2	3.9	organelle organization and biogenesis
YLR292C	SEC72	49	1	2	3.9	transport
YKL025C	PAN3	49	1	2	3.9	DNA metabolism
YDR183W	PLP1	49	1	2	3.9	cytoskeleton organization and biogenesis
YOR026W	BUB3	47	3	6	3.8	cell cycle
YPR135W	CTF4	49	1	2	3.7	DNA metabolism
YBR103W	SIF2	48	2	4	3.7	meiosis
YFR019W	FAB1	49	1	2	3.7	response to stress
YGR188C	BUB1	47	3	6	3.6	protein modification
YDR485C	VPS72	41	9	18	3.3	transport
YPL253C	VIK1	50	0	0	3.1	cytoskeleton organization and biogenesis
YNL140C		50	0	0	2.8	unknown

Table S3. Null mutant of previously uncharacterized yeast ORF *YLL049W* exhibits temperature-dependent nuclear migration defect similar to Dynactin component *JNM1* and distinct from the temperature-independent defect of Kinesin-related gene *KIP2*.

	Mutant	Normal	Abnormal
30 ºC	BY4741 (WT)	97	3
	jnm1Δ	86	14
	kip2Δ	87	13
	yll049wΔ	85	15
13 ºC	BY4741 (WT)	94	6
	jnm1Δ	38	62
	kip2Δ	86	14
	yll049wΔ	42	58

Table S4. Benomyl sensitivity at 5 µg/ml for mutants congruent to CIN1. Approximately equal amounts (OD_{600}) of each mutant were arrayed with three five-fold serial dilutions on media with and without 5 µg/ml Benomyl in DMSO using a 96-pin transfer device. Mutants were blind scored as Benomyl sensitive if they displayed any decrease in growth compared to DMSO alone.

ORF	Name	Score	Ben[S]	GO slim Biological Process	DMSO Control Cells: ◤	5 µg/ml Benomyl Cells: ◤
(BY4741)	(WT)	NA	No			
YOR349W	CIN1	NA	Yes	cytoskeleton organization and biogenesis		
YER007W	PAC2	14.63	Yes	cytoskeleton organization and biogenesis		
YPL241C	CIN2	13.41	Yes	cytoskeleton organization and biogenesis		
YEL061C	CIN8	12.40	Yes	cell cycle		
YML124C	TUB3	9.49	Yes	meiosis		
YPL269W	KAR9	9.06	Yes	nuclear organization and biogenesis		
YEL003W	GIM4	8.74	No	cytoskeleton organization and biogenesis		
YER016W	BIM1	8.68	No	cytoskeleton organization and biogenesis		
YMR138W	CIN4	7.77	Yes	cytoskeleton organization and biogenesis		
YLR200W	YKE2	6.98	Yes	cytoskeleton organization and biogenesis		
YGL086W	MAD1	6.33	No	transport		
YOR265W	RBL2	6.06	No	protein binding		

ORF	Name	Score	Ben^S	GO slim Biological Process	DMSO Control Cells:	5 μg/ml Benomyl Cells:
YGL217C		6.06	No	unknown		
YJR053W	BFA1	5.96	No	conjugation		
YLR210W	CLB4	5.28	No	cell cycle		
YCL029C	BIK1	5.28	Yes	cell cycle		
YJL013C	MAD3	5.28	No	cell cycle		
YDR360W		5.24	No	unknown		
YOR058C	ASE1	4.74	No	cell cycle		
YPL008W	CHL1	4.57	No	cell cycle		
YMR055C	BUB2	4.57	No	cell cycle		
YGL216W	KIP3	4.57	Yes	cytoskeleton organization and biogenesis		
YBR107C	IML3	4.53	No	cell cycle		
YJL030W	MAD2	4.52	No	cell cycle		
YOR264W	DSE3	4.44	No	unknown		
YOR266W	PNT1	4.44	No	membrane organization and biogenesis		
YPL017C		4.40	No	unknown		
YGL124C	MON1	4.40	No	transport		
YOR026W	BUB3	4.28	Yes	cell cycle		
YLR381W	CTF3	4.03	No	cell cycle		
YJR135C	MCM22	4.03	No	cell cycle		

Table S5. The list of genes having average congruence score ≥ 4 with 7 benomyl

sensitive landmarks.

ORF name	Gene name	LD$_{50}$ benomyl concentration (μg/ml)	Average congruence score with 7 landmarks	GO slim Biological Process
YML124C	TUB3	1	4	meiosis
YGR188C	BUB1	15	4	protein modification
YJL030W	MAD2	20	4	cell cycle
YCL029C	BIK1	20	4	cell cycle
YPL241C	CIN2	5	5	cytoskeleton organization and biogenesis
YGL086W	MAD1	20	5	transport
YOR349W	CIN1	1	6	cytoskeleton organization and biogenesis
YLR200W	YKE2	5	8	cytoskeleton organization and biogenesis
YGR078C	PAC10	1	10	cytoskeleton organization and biogenesis
YNL153C	GIM3	1	11	cytoskeleton organization and biogenesis
YML094W	GIM5	1	12	cytoskeleton organization and biogenesis
YEL003W	GIM4	5	12	cytoskeleton organization and biogenesis

Table S6. The list of genes having ≥ 4 synthetic lethal interactions with 7 benomyl-sensitive landmarks.

ORF name	Gene name	LD$_{50}$ benomyl concentration (μg/ml)	Number of synthetic lethal interactions with 7 landmarks	GO slim Biological Process
YML124C	TUB3	1	4	meiosis
YOR349W	CIN1	1	4	cytoskeleton organization and biogenesis
YAL011W	SWC1	15	4	organelle organization and biogenesis
YJL030W	MAD2	20	4	cell cycle
YDR318W	MCM21	30	4	cell cycle
YCL016C	DCC1	30	4	cell cycle
YPL018W	CTF19	30	4	cell cycle
YHR191C	CTF8	35	4	cell cycle
YJL013C	MAD3	1	5	cell cycle
YPL241C	CIN2	5	5	cytoskeleton organization and biogenesis
YMR138W	CIN4	10	5	cytoskeleton organization and biogenesis
YOR265W	RBL2	20	5	protein binding
YGL086W	MAD1	20	5	transport
YCL029C	BIK1	20	5	cell cycle
YOL012C	HTZ1	20	5	transcription
YLR085C	ARP6	20	5	transport

Table S7. *PFD1* dSLAM targets. Genes exhibiting synthetic lethality in combination with *PFD1* were detected by microarray analysis after sporulation of a heterozygous deletion mutant pool that had been transformed with a *pfd1* knockout allele. The experimental (*pfd1 yko*): control (*yko*) tag signal ratios were determined from Uptag and Downtag hybridizations. In our experience, a signal ratio > 2 for both tags represents a conservative criterion for identification of true synthetic lethal relationships (minimizing false positives). Protein complex residence of *PFD1* dSLAM targets is indicated by the name of the bait protein used to identify complex members.

ORF Name	Gene Name	log$_2$(Ratio) Downtag	log$_2$(Ratio) Uptag	Protein Complex Residence
YMR074C		4.71	3.99	
YOR349W	CIN1	5.35	3.93	CDC55[Ho et al. 2002]
YDR334W	SWR1	3.7	5.12	
YML124C	TUB3	3.53	4.69	HIS4[Gavin et al. 2002], RAD3[Gavin et al. 2002], YDR060W[Gavin et al. 2002], YLL013C[Gavin et al. 2002]
YPL241C	CIN2	3.59	3.19	
YOL012C	HTZ1	3.05	3.66	NAP1[Gavin et al. 2002], RAD16[Ho et al. 2002]
YLR085C	ARP6	2.92	3.33	
YKL025C	PAN3	2.65	2	PAN2[Gavin et al. 2002]
YNL054W	VAC7	1.91	2.48	
YOR073W	SGO1	1.81	2.85	BUD32[Ho et al. 2002]
YCL029C	BIK1	2.22	1.73	LAP4[Ho et al. 2002]
YML112W	CTK3	1.72	2.61	CTK1[Gavin et al. 2002], CTK3[Ho et al. 2002]
YJL030W	MAD2	1.69	1.55	
YKL037W		1.31	1.6	
YCR009C	RVS161	1.29	1.8	RVS161[Ho et al. 2002], RVS167[Ho et al. 2002], SEC27[Ho et al. 2002]
YBR231C	AOR1	1.24	3.47	
YPL174C	NIP100	1.22	1.34	
YNL148C	ALF1	1.21	2.88	
YER016W	BIM1	1.46	1.2	
YJL129C	TRK1	1.2	1.25	
YNL248C	RPA49	1.17	1.45	RPA190[Gavin et al. 2002], RPC40[Gavin et al. 2002],[Ho et al. 2002]
YNL296W	KRE25	1.41	1.17	
YNL273W	TOF1	1.16	1.26	

ORF Name	Gene Name	log$_2$(Ratio) Downtag	log$_2$(Ratio) Uptag	Protein Complex Residence
YER177W	BMH1	1.39	1.16	BMH1[Ho et al. 2002], BMH2[Gavin et al. 2002], LCB2[Gavin et al. 2002], SNF4[Gavin et al. 2002]
YLR370C	ARC18	1.25	1.14	ARC18[Gavin et al. 2002], ARC40[Ho et al. 2002], ARP2[Ho et al. 2002]
YBR036C	CSG2	1.33	1.13	
YCR024C		1.13	1.48	
YLR442C	SIR3	1.08	1.59	
YGL094C	PAN2	1.07	2.55	PAN2[Gavin et al. 2002]
YER087W		1.2	1.03	
YNL086W		1.28	1.02	
YDL020C	RPN4	1.02	1.46	
YPR141C	KAR3	1.05	1.01	
YDR207C	UME6	1.01	1.76	YDL076C[Gavin et al. 2002]

Table S8. Prefoldin Congruence SGA-SGA and dSLAM-SGA: i, Gene 1 SL interaction set size; j, Gene 2 interaction set size; k, interaction set overlap; *Score*, $-\text{Log}_{10}(P)$. The total number of target genes is 4700. Because we used conservative dSLAM criteria to identify interactions, only those mutants scored as synthetic lethal (not synthetic sick) from the SGA data were kept for the congruence score comparison.

Comparison	Gene 1	Gene 2	i	j	k	Score
SGA-SGA	GIM3	GIM4	66	50	36	60
	GIM3	GIM5	66	29	19	29
	GIM3	PAC10	66	72	43	67
	GIM3	YKE2	66	47	36	62
	GIM4	GIM5	50	29	16	25
	GIM4	PAC10	50	72	30	44
	GIM4	YKE2	50	47	26	42
	GIM5	PAC10	29	72	19	28
	GIM5	YKE2	29	47	15	23
	PAC10	YKE2	72	47	34	55
dSLAM-SGA	PFD1	GIM3	33	66	12	14
	PFD1	GIM4	33	50	11	14
	PFD1	GIM5	33	29	7	9
	PFD1	PAC10	33	72	13	15
	PFD1	YKE2	33	47	12	16

Table S9. The dSLAM screen results for query genes *LTE1*, *SPO12*, and *SLK19*.

Every synthetic lethal interaction has been confirmed by either random spore analysis

(RSA) or tetrad analysis.

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YAL013W	DEP1	SL	
YAL024C	LTE1	YAR003W	SWD1	SF	
YAL024C	LTE1	YBL016W	FUS3	SF	
YAL024C	LTE1	YBL025W	RRN10	SF	
YAL024C	LTE1	YBL031W	SHE1	SF	
YAL024C	LTE1	YBL032W	HEK2	SF	
YAL024C	LTE1	YBL058W	SHP1	SL	
YAL024C	LTE1	YBR036C	CSG2	SF	
YAL024C	LTE1	YBR058C	UBP14	SF/SL	
YAL024C	LTE1	YBR097W	VPS15	SF/SL	
YAL024C	LTE1	YBR119W	MUD1	SF/SL	
YAL024C	LTE1	YBR174C		SF	
YAL024C	LTE1	YBR175W	SWD3	SF	
YAL024C	LTE1	YBR200W	BEM1	SL	
YAL024C	LTE1	YBR267W		SF	
YAL024C	LTE1	YCL016C	DCC1	SL	
YAL024C	LTE1	YCL037C	SRO9	SF	
YAL024C	LTE1	YCL060C		SF	
YAL024C	LTE1	YCL061C	MRC1	SF	
YAL024C	LTE1	YCL063W	VAC17	SF	
YAL024C	LTE1	YCR016W		SF	
YAL024C	LTE1	YCR066W	RAD18	SF	
YAL024C	LTE1	YCR094W	CDC50	SF	
YAL024C	LTE1	YDL006W	PTC1	SF	
YAL024C	LTE1	YDL040C	NAT1	SF	
YAL024C	LTE1	YDL056W	MBP1	SF	
YAL024C	LTE1	YDL059C	RAD59	SF	
YAL024C	LTE1	YDL074C	BRE1	SF/SL	
YAL024C	LTE1	YDL090C	RAM1	SF	
YAL024C	LTE1	YDL115C	IWR1	SL	
YAL024C	LTE1	YDL130W	RPP1B	SF	
YAL024C	LTE1	YDL136W	RPL35B	SF	
YAL024C	LTE1	YDL144C		SF	
YAL024C	LTE1	YDL190C	UFD2	SF	
YAL024C	LTE1	YDL225W	SHS1	SF	

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YDL236W	PHO13	SF	
YAL024C	LTE1	YDR004W	RAD57	SF	
YAL024C	LTE1	YDR065W		SF	
YAL024C	LTE1	YDR071C		SF	
YAL024C	LTE1	YDR076W	RAD55	SF	
YAL024C	LTE1	YDR101C	ARX1	SF	
YAL024C	LTE1	YDR114C		SF	
YAL024C	LTE1	YDR117C		SF	
YAL024C	LTE1	YDR121W	DPB4	SF	
YAL024C	LTE1	YDR146C	SWI5	SF	
YAL024C	LTE1	YDR149C		SF	
YAL024C	LTE1	YDR150W	NUM1	SF	
YAL024C	LTE1	YDR159W	SAC3	SL	
YAL024C	LTE1	YDR174W	HMO1	SF	
YAL024C	LTE1	YDR200C	VPS64	SF	
YAL024C	LTE1	YDR207C	UME6	SL	
YAL024C	LTE1	YDR260C	SWM1	SF/SL	
YAL024C	LTE1	YDR310C	SUM1	SL	
YAL024C	LTE1	YDR359C	VID21	SF	
YAL024C	LTE1	YDR369C	XRS2	SF	
YAL024C	LTE1	YDR392W	SPT3	SF/SL	
YAL024C	LTE1	YDR432W	NPL3	SL	
YAL024C	LTE1	YDR463W	STP1	SF	
YAL024C	LTE1	YDR469W	SDC1	SF	
YAL024C	LTE1	YDR497C	ITR1	SF/SL	
YAL024C	LTE1	YDR532C		SL	
YAL024C	LTE1	YEL029C	BUD16	SF	
YAL024C	LTE1	YEL031W	SPF1	SF	
YAL024C	LTE1	YEL037C	RAD23	SF	
YAL024C	LTE1	YEL054C	RPL12A	SF	
YAL024C	LTE1	YEL061C	CIN8	SF	
YAL024C	LTE1	YEL062W	NPR2	SF	
YAL024C	LTE1	YER014W	HEM14	SF	
YAL024C	LTE1	YER016W	BIM	SL	
YAL024C	LTE1	YER073W	ALD5	SL	
YAL024C	LTE1	YER095W	RAD51	SF	
YAL024C	LTE1	YER110C	KAP123	SF	
YAL024C	LTE1	YER123W	YCK3	SF	
YAL024C	LTE1	YER139C		SF	
YAL024C	LTE1	YFR036W	CDC26	SF	
YAL024C	LTE1	YGL045W	RIM8	SF	

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YGL060W		SF	
YAL024C	LTE1	YGL066W	SGF73	SL	
YAL024C	LTE1	YGL072C		SL	
YAL024C	LTE1	YGL078C	DBP3	SF	
YAL024C	LTE1	YGL127C	SOH1	SF/SL	
YAL024C	LTE1	YGL133W	ITC1	SF/SL	
YAL024C	LTE1	YGL163C	RAD54	SF	
YAL024C	LTE1	YGL167C	PMR1	SF	
YAL024C	LTE1	YGL228W	SHE10	SF	
YAL024C	LTE1	YGR046W		SF	
YAL024C	LTE1	YGR056W	RSC1	SF	
YAL024C	LTE1	YGR077C	PEX8	SF	
YAL024C	LTE1	YGR078C	PAC10	SF	
YAL024C	LTE1	YGR134W	CAF130	SF	
YAL024C	LTE1	YGR180C	RNR4	SL	
YAL024C	LTE1	YGR192C	TDH3	SF/SL	
YAL024C	LTE1	YGR260W	TNA1	SF	
YAL024C	LTE1	YHL007C	STE20	SL	
YAL024C	LTE1	YHL027W	RIM101	SF	
YAL024C	LTE1	YHL033C	RPL8A	SF/SL	
YAL024C	LTE1	YHR013C	ARD1	SF	
YAL024C	LTE1	YHR031C	RRM3	SF	
YAL024C	LTE1	YHR034C		SF	
YAL024C	LTE1	YHR041C	SRB2	SL	
YAL024C	LTE1	YHR067W	RMD12	SF	
YAL024C	LTE1	YHR100C		SF	
YAL024C	LTE1	YHR129C	ARP1	SF	
YAL024C	LTE1	YHR152W	SPO12	SL	
YAL024C	LTE1	YHR154W	RTT107	SF	
YAL024C	LTE1	YHR178W	STB5	SL	
YAL024C	LTE1	YHR191C	CTF8	SL	
YAL024C	LTE1	YHR200W	RPN10	SL	
YAL024C	LTE1	YIL036W	CST6	SF	
YAL024C	LTE1	YIL084C	SDS3	SL	
YAL024C	LTE1	YIL103W		SF	
YAL024C	LTE1	YIR023W	DAL81	SF	
YAL024C	LTE1	YIR033W	MGA2	SF	
YAL024C	LTE1	YJL047C	RTT101	SF	
YAL024C	LTE1	YJL080C	SCP160	SL	
YAL024C	LTE1	YJL098W	SAP185	SF	
YAL024C	LTE1	YJL115W	ASF1	SF/SL	

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YJL120W		SF	
YAL024C	LTE1	YJL121C	RPE1	SF	
YAL024C	LTE1	YJL128C	PBS2	SF	
YAL024C	LTE1	YJL148W	RPA34	SF	
YAL024C	LTE1	YJL177W	RPL17B	SF	
YAL024C	LTE1	YJL179W	PFD1	SF	
YAL024C	LTE1	YJR043C	POL32	SF	
YAL024C	LTE1	YJR050W	ISY1	SF	
YAL024C	LTE1	YJR055W	HIT1	SL	
YAL024C	LTE1	YJR055W	HIT1	SF	
YAL024C	LTE1	YJR074W	MOG1	SL	
YAL024C	LTE1	YJR097W		SF	
YAL024C	LTE1	YJR102C	VPS25	SF	
YAL024C	LTE1	YKL006W	RPL14A	SL	
YAL024C	LTE1	YKL053W		SF	
YAL024C	LTE1	YKL074C	MUD2	SF	
YAL024C	LTE1	YKL113C	RAD27	SF	
YAL024C	LTE1	YKR047W		SF	
YAL024C	LTE1	YKR048C	NAP1	SF	
YAL024C	LTE1	YKR054C	DYN1	SF	
YAL024C	LTE1	YKR061W	KTR2	SL	
YAL024C	LTE1	YKR073C		SF	
YAL024C	LTE1	YKR092C	SRP40	SF	
YAL024C	LTE1	YLL002W	RTT109	SF	
YAL024C	LTE1	YLL049W		SF	
YAL024C	LTE1	YLR015W	BRE2	SF	
YAL024C	LTE1	YLR027C	AAT2	SL	
YAL024C	LTE1	YLR032W	RAD5	SF	
YAL024C	LTE1	YLR055C	SPT8	SL	
YAL024C	LTE1	YLR067C	PET309	SF	
YAL024C	LTE1	YLR079W	SIC1	SF	
YAL024C	LTE1	YLR102C	APC9	SL	
YAL024C	LTE1	YLR200W	YKE2	SF	
YAL024C	LTE1	YLR204W	QRI5	SF	
YAL024C	LTE1	YLR234W	TOP3	SL	
YAL024C	LTE1	YLR235C		SL	
YAL024C	LTE1	YLR240W	VPS34	SF	
YAL024C	LTE1	YLR315W	NKP2	SF	
YAL024C	LTE1	YLR320W	MMS22	SL	
YAL024C	LTE1	YLR338W		SL	
YAL024C	LTE1	YLR357W	RSC2	SL	

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YLR358C		SL	
YAL024C	LTE1	YLR370C	ARC18	SL	
YAL024C	LTE1	YLR373C	VID22	SF	
YAL024C	LTE1	YLR374C		SF	
YAL024C	LTE1	YLR386W	VAC14	SF	
YAL024C	LTE1	YLR406C	RPL31B	SF	
YAL024C	LTE1	YLR410W	VIP1	SL	
YAL024C	LTE1	YLR417W	VPS36	SF	
YAL024C	LTE1	YLR448W	RPL6B	SF/SL	
YAL024C	LTE1	YML032C	RAD52	SF	
YAL024C	LTE1	YML036W		SL	
YAL024C	LTE1	YML061C	PIF1	SF	
YAL024C	LTE1	YML094W	GIM5	SF	
YAL024C	LTE1	YML103C	NUP188	SF	
YAL024C	LTE1	YML128C	MSC1	SF/SL	
YAL024C	LTE1	YMR022W	QRI8	SF	
YAL024C	LTE1	YMR039C	SUB1	SF	
YAL024C	LTE1	YMR048W	CSM3	SF	
YAL024C	LTE1	YMR063W	RIM9	SF	
YAL024C	LTE1	YMR078C	CTF18	SL	
YAL024C	LTE1	YMR144W		SF	
YAL024C	LTE1	YMR154C	RIM13	SF	
YAL024C	LTE1	YMR165C	SMP2	SL	
YAL024C	LTE1	YMR179W	SPT21	SL	
YAL024C	LTE1	YMR194W	RPL36A	SF	
YAL024C	LTE1	YMR198W	CIK1	SF/SL	
YAL024C	LTE1	YMR198W	CIK1	SL	
YAL024C	LTE1	YMR205C	PFK2	SL	
YAL024C	LTE1	YMR214W	SCJ1	SF	
YAL024C	LTE1	YMR224C	MRE11	SF	
YAL024C	LTE1	YMR261C	TPS3	SF	
YAL024C	LTE1	YMR263W	SAP30	SL	
YAL024C	LTE1	YMR267W	PPA2	SF	
YAL024C	LTE1	YMR269W		SF	
YAL024C	LTE1	YMR274C	RCE1	SF	
YAL024C	LTE1	YMR294W	JNM1	SF	
YAL024C	LTE1	YMR299C		SF	
YAL024C	LTE1	YNL054W	VAC7	SF/SL	
YAL024C	LTE1	YNL064C	YDJ1	SL	
YAL024C	LTE1	YNL068C	FKH2	SF	
YAL024C	LTE1	YNL076W	MKS1	SF/SL	

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YNL084C	END3	SF/SL	
YAL024C	LTE1	YNL097C	PHO23	SL	
YAL024C	LTE1	YNL147W	LSM7	SF/SL	
YAL024C	LTE1	YNL148C	ALF1	SF	
YAL024C	LTE1	YNL153C	GIM3	SF	
YAL024C	LTE1	YNL171C		SL	
YAL024C	LTE1	YNL198C		SF	
YAL024C	LTE1	YNL199C	GCR2	SL	
YAL024C	LTE1	YNL229C	URE2	SF	
YAL024C	LTE1	YNL236W	SIN4	SF	
YAL024C	LTE1	YNL250W	RAD50	SF	
YAL024C	LTE1	YNL273W	TOF1	SF	
YAL024C	LTE1	YNL294C	RIM21	SL	
YAL024C	LTE1	YNL330C	RPD3	SL	
YAL024C	LTE1	YNR009W		SF	
YAL024C	LTE1	YOL004W	SIN3	SL	
YAL024C	LTE1	YOL041C	NOP12	SF/SL	
YAL024C	LTE1	YOL068C	HST1	SF	
YAL024C	LTE1	YOR035C	SHE4	SF	
YAL024C	LTE1	YOR080W	DIA2	SL	
YAL024C	LTE1	YOR082C		SF	
YAL024C	LTE1	YOR083W	WHI5	SF	
YAL024C	LTE1	YOR195W	SLK19	SL	
YAL024C	LTE1	YOR209C	NPT1	SL	
YAL024C	LTE1	YOR211C	MGM1	SL	
YAL024C	LTE1	YOR221C	MCT1	SF	
YAL024C	LTE1	YOR271C		SF	
YAL024C	LTE1	YOR275C	RIM20	SF	
YAL024C	LTE1	YOR295W	UAF30	SL	
YAL024C	LTE1	YOR297C	TIM18	SF	
YAL024C	LTE1	YOR304W	ISW2	SL	
YAL024C	LTE1	YOR344C	TYE7	SF	
YAL024C	LTE1	YOR360C	PDE2	SF	
YAL024C	LTE1	YPL008W	CHL1	SF	
YAL024C	LTE1	YPL055C	LGE1	SF	
YAL024C	LTE1	YPL059W	GRX5	SF	
YAL024C	LTE1	YPL080C		SF	
YAL024C	LTE1	YPL084W	BRO1	SL	
YAL024C	LTE1	YPL106C	SSE1	SF	
YAL024C	LTE1	YPL139C	UME1	SF	
YAL024C	LTE1	YPL161C	BEM4	SL	

Query ORF	Query Gene	Target ORF	Target Gene	RSA	TETRAD
YAL024C	LTE1	YPL174C	NIP100	SF	
YAL024C	LTE1	YPL178W	CBC2	SF	
YAL024C	LTE1	YPL182C		SF	
YAL024C	LTE1	YPL184C		SF	
YAL024C	LTE1	YPL188W	POS5	SL	
YAL024C	LTE1	YPL213W	LEA1	SF	
YAL024C	LTE1	YPL269W	KAR9	SF	
YAL024C	LTE1	YPR029C	APL4	SF	
YAL024C	LTE1	YPR054W	SMK1	SF	
YAL024C	LTE1	YPR119W	CLB2	SF	
YAL024C	LTE1	YPR135W	CTF4	SL	
YAL024C	LTE1	YPR141C	KAR3	SF	
YHR152W	SPO12	YAL024C	LTE1		SL
YHR152W	SPO12	YGL003C	CDH1		SL
YHR152W	SPO12	YNL171C			SL
YHR152W	SPO12	YNL225C	CNM67		SL
YHR152W	SPO12	YNL298W	CLA4		SL
YOR195W	SLK19	YAL024C	LTE1		SL
YOR195W	SLK19	YCR086W	CSM1		SF
YOR195W	SLK19	YDR200C	VPS64		SL
YOR195W	SLK19	YDR359C	VID21		SL
YOR195W	SLK19	YDR439W	LRS4		SL
YOR195W	SLK19	YEL061C	CIN8		SL
YOR195W	SLK19	YER016W	BIM1		SL
YOR195W	SLK19	YGL003C	CDH1		SL
YOR195W	SLK19	YGR188C	BUB1		SL
YOR195W	SLK19	YHR191C	CTF8		SF
YOR195W	SLK19	YJL124C	LSM1		SL
YOR195W	SLK19	YKL057C	NUP120		SL
YOR195W	SLK19	YKR082W	NUP133		SL
YOR195W	SLK19	YML112W	CTK3		SL
YOR195W	SLK19	YMR078C	CTF18		SL
YOR195W	SLK19	YMR198W	CIK1		SL
YOR195W	SLK19	YNL225C	CNM67		SL
YOR195W	SLK19	YNL298W	CLA4		SL
YOR195W	SLK19	YOR026W	BUB3		SL

Table S10. Congruence scores for FEAR and MEN pathway members: i, Gene 1 SL

interaction set size; j, Gene 2 interaction set size; k, interaction set overlap; *Score*,

$-\text{Log}_{10}$(hypergeometric *P*-value). The total number of target genes is 4700.

Comparison	Gene 1	Gene 2	i	j	k	Score
	SPO12	SLK19	5	19	4	9
dSLAM-dSLAM	SPO12	LTE1	5	252	1	1
	SLK19	LTE1	19	252	7	4
	CLA4	LTE1	67	252	31	22
dSLAM-SGA	CLA4	SPO12	67	5	0	0
	CLA4	SLK19	67	19	2	2

Chapter 6.

Predicting pathways in yeast using genome-wide phenotype data.

Introduction

Deletion of genes operating in the same pathway results in a similar synthetic lethal interaction profile (Ye and Peyser et al. 2005). A measurement termed the "congruence score" describes the similarity of genetic interaction sets for each pair of mutants. High scores are associated with genes exhibiting similar genetic interactions, which function in the same pathway.

The relative growth rates of all yeast knockout (YKO) strains in the presence of a second allele (the "query") can be thought of as a complex set of phenotypes. It is possible to generalize this to other phenotypes beyond growth in presence versus absence of a query allele. As with genetic congruence (Ye and Peyser et al. 2005), mutants that share similar phenotypes are likely to be components of the same pathway. For example, growth in presence of benomyl (Pan et al. 2004), or in various media (as in Giaever et al. 2002). In Lee and St Onge et al. (2005), mutants were grouped by sensitivity to 12 DNA-damaging agents using hierarchical clustering. Two-dimensional hierarchical clustering is useful for aligning mutants and treatments by similarity, but does not perform well when the source data distributions are varied.

Also included in Giaever et al. (2002) were cell morphology phenotypes. Phenotypes such as 'round' or 'elongated' are not well-incorporated into hierarchical clustering techniques typically relying on Pearson or Spearman correlation coefficients. However, it is possible to generate a similarity score for each pair of deletion mutants by adapting Resnik's (1995) application of information entropy (Shannon 1948) to shared information content. This shared information content score is based on intersections

191

within a directed acyclic graph called an ontology. Lord et al. (2003) applied this concept to the Gene Ontology (GO) (Ashburner et al. 2000) to describe the similarity of annotations for human genes. By translating the shared information technique to distributions of phenotype data, I was able to generate a phenotype similarity score. This approach does not require similar data and scores are readily updated. I collected 200 genome-wide yeast phenotype sets from published reports and applied the shared information calculation to all pairs of mutants across all phenotypes.

Methods

Data were collected from published reports of phenotypes for YKOs (Table 1). Only screens or selections using an entire collection were obtained. Phenotypes for 5918 open reading frames (ORFs) mutated in each strain were assigned a numeric value, such as 1 or 0 for YKOs that are round or not round, respectively. Phenotypes with a severity were assigned appropriately increasing numeric values, and phenotypes with numeric values were rounded to 1 decimal place and used directly. Data were loaded into R (Ihaka and Gentleman 1996), and collected into a set of 5918 numbers corresponding to a list of deleted ORFs. Each list of 5918 values was then written to a single text file with one value per line.

The group of 200 phenotype sets were analyzed using perl. First, the list of ORFs corresponding to each value was read, then for each phenotype the list of values was read. For each pair of ORFs and phenotype, the phenotype similarity score was calculated as the negative logarithm of the fraction of data contained within the inclusive interval between the two values. This is equal to $-\log_{10}(P)$, where P is the chance any randomly

chosen ORF will fall within the range of values defined by the first and second ORFs (see

Figure 1). Then for each successive phenotype, the scores for each pair were added,

resulting in a final phenotype similarity score of

$$\sum_{i=1}^{n} -\log_{10}(P_i) = -\log_{10}\left(\prod_{i=1}^{n} P_i\right) \ ,$$

where n = number of phenotypes and P_i is the fraction of values found in the inclusive

interval between the ORF values, for each phenotype. This measure can be expressed in

terms of bits of information if $\log_2$ is used in place of $\log_{10}$. Using $\log_{10}$ results in a

measure of decimal digits rather than binary digits, but a decimal digit is equivalent to

3.322 bits (log 10/log 2). I express similarity in decimal format here.

Similarity of GO annotations for each gene pair was calculated using a method after

Lord, et al. (2003). Briefly, the annotation similarity is expressed as $-\log_2(P)$, where P is

the chance of randomly selecting an annotation that is a term shared by both ORFs. The

maximum value for similarity is assigned to the pair. There are 3 ontologies within GO:

Biological Process, Cellular Component, and Molecular Function. For each of these

ontologies, the root term is the name of the ontology, and a gene's annotation consists of

the directed acyclic graph from the root to the most specific term. Genes without

annotation for biological process, for example, are listed as: "biological_process" →

"biological_process unknown," whereas a gene involved in cell-cell adhesion would be

annotated "biological_process" → "biological adhesion" → "cell adhesion" → "cell-cell

adhesion." A gene annotated "biofilm formation" would share "cell adhesion" as the

most specific annotation with a gene annotated "cell-cell adhesion" and the gene pair

would be assigned the Biological Process similarity value of $-\log_2(17/5331) = 8.3$, where 17 is the number of "cell-cell adhesion" annotations and 5331 is the number of "biological_process" annotations. GO annotations were obtained from *Saccharomyces* Genome Database (http://www.yeastgenome.org) as "gene_association.sgd" (ftp://genome-ftp.stanford.edu/pub/yeast/literature_curation/gene_association.sgd) on 2006 December 15, version 1.1308.

Protein-protein interactions were taken from the Krogan et al. (2006) "Core Network" interaction set.

Results

I tested the phenotype similarity score for its ability to connect functionally related genes by comparing the similarity of GO annotations for increasing similarity of phenotypes. With increasing cutoffs for phenotypic similarity, the decreasing number of gene pairs are enriched for similar GO annotation (Figure 2). Biological Process and Cellular Component display higher similarity than Molecular Function. This is expected, since similar phenotypes should be more closely associated with genes that function in the same pathway, than with genes that carry out similar molecular reactions.

When gene products function in the same pathway, they are more likely to physically interact. Therefore, phenotypically similar genes should encode proteins that are more likely to interact. I calculated the fraction of gene pairs above various phenotype similarity score cutoffs that encode proteins for which physical interaction has been reported in a high-throughput protein-protein interaction study (Krogan et al. 2006). The

194

fraction of interacting pairs increases with phenotypic similarity, and the number of

interacting pairs is higher than expected by chance ($P \leq 0.05$) above phenotypic similarity

of 69, with P calculated using a one-sided Fisher's exact test (Figure 3). At phenotype

similarity score of 100 or more, $P = 7 \times 10^{-30}$.

With ~10^3 gene pairs above phenotype similarity score of 100, that was chosen as a

cutoff to display a network connecting similar genes (Figure 4). The entire network at

≥ 100 consists of 697 nodes representing genes and 1534 edges representing phenotypic

similarity. The largest connected component is 537 nodes and 1418 edges. At phenotype

similarity score ≥ 95, there are 5106 interactions among 1289 genes, and at ≥ 90, there

are 16 904 interactions among 2282 genes. Many genes known to perform related

functions are connected, and novel associations are suggested. Examples of

interconnected subsets are: genes related to DNA damage (Figure 5), and genes related to

ribosome structure and function (Figure 6). Along with many structural ribosomal genes

(*RPS#* and *RPL#*), the ribosome subset includes *ARC1*, which ensures tRNA delivery to

the cytoplasm (Galani et al. 2001) and *TSR2*, which is involved in 20S pre-rRNA

processing (Peng et al. 2003). Figure 7 shows a small sub-network containing genes

related to biosynthesis of tryptophan. All connected genes are members of the

superpathway of phenylalanine, tyrosine and tryptophan biosynthesis upstream of or

within the tryptophan biosynthesis pathway. The complete list of genes in this segment of

the pathway is: *ARO1*, *ARO2*, *ARO3*, *ARO4*, *TRP1*, *TRP2*, *TRP3*, *TRP4*, and *TRP5*. At a

more relaxed phenotypic similarity cutoff of 90, all these tryptophan pathway genes are

connected with the exception of *ARO4*, which displays no interactions at phenotypic

similarity of 90 or higher.

Discussion

Introduction of the YKO collection has allowed genome-wide examination of phenotypes in deletion mutants. Similar phenotypes are expected for genes that function in a single pathway, and I present a method for quantifying that similarity. This information-content–based approach overcomes some of the problems with current methods, and is able to handle any measure of phenotype. Another pleasing aspect of this method is that when new data become available, they can be readily combined. Each gene pair is given a similarity score for that new phenotype, and those scores are simply added to the previously existing totals.

Results from this calculation are complex and visualization in a graph (Figure 4) presents a large problem. It is likely that, as with high-throughput protein interaction studies, the graph contains interactions that do not represent informative relationships. Simplifying the network by assignment of genes to interconnected modules responsible for some pathway could be accomplished with algorithms such as Markov Clustering (Enright et al. 2002; Brohee and van Helden 2006). Alternatively, subsections of the graph corresponding to genes of interest can be viewed, such as in Figures 5, 6, and 7.

An important use for these relationships is as verification of functionally important physical interactions. Combination of these data with reported protein-protein interaction sets can refine imperfect information. Overlapping indications of interaction by genetic and physical evidence strongly argue for shared pathway membership. Additionally,

phenotypic similarity can group genes into pathways when they do not physically

interact. An example is provided in Figure 7, with genes involved in biosynthesis of

tryptophan connected. Among the genes in this pathway, only Trp2p-Trp3p and

Trp3p-Aro1p have been reported to physically interact (Gavin et al. 2006). The other

genes are connected by small molecules that they pass along in the pathway. The ability

of phenotypic similarity to find these connections provides additional value over

networks based only on physical interactions.

Acknowledgments

References

Aboussekhra A, Chanet R, Zgaga Z, Cassier-Chauvat C, Heude M, Fabre F. 1989. RADH, a gene of *Saccharomyces cerevisiae* encoding a putative DNA helicase involved in DNA repair. Characteristics of radH mutants and sequence of the gene. *Nucleic Acids Res* 17(18):7211–9.

Alcasabas AA, Osborn AJ, Bachant J, Hu F, Werler PJ, Bousset K, Furuya K, Diffley JF, Carr AM, Elledge SJ. 2001. Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nat Cell Biol* 3(11):958–65.

Ando A, Tanaka F, Murata Y, Takagi H, Shima J. 2006. Identification and classification of genes required for tolerance to high-sucrose stress revealed by genome-wide screening of *Saccharomyces cerevisiae*. *FEMS Yeast Res* 6(2):249–67.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP,

Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25–9.

Blackburn AS, Avery SV. 2003. Genome-wide screening of *Saccharomyces cerevisiae* to identify genes required for antibiotic insusceptibility of eukaryotes. *Antimicrob Agents and Chemotherapy* 47(2):676–81.

Brohee S, van Helden J. 2006. Evaluation of clustering algorithms for protein-protein interaction networks. *BMC Bioinformatics* 7:488.

Corbacho I, Olivero I, Hernandez LM. 2005. A genome-wide screen for *Saccharomyces cerevisiae* nonessential genes involved in mannosyl phosphate transfer to mannoprotein-linked oligosaccharides. *Fung Genet Biol* 42(9):773–90.

Dorer RK, Zhong S, Tallarico JA, Wong WH, Mitchison TJ, Murray AW. 2005. A small-molecule inhibitor of Mps1 blocks the spindle-checkpoint response to a lack of tension on mitotic chromosomes. *Current Biol* 15(11):1070–6.

Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30(7):1575–84.

Galani K, Grosshans H, Deinert K, Hurt EC, Simos G. 2001. The intracellular location of two aminoacyl-tRNA synthetases depends on complex formation with Arc1p. *EMBO J* 20(23):6889–98.

Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B, et al. 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440(7084):631–6.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418(6896):387–91.

Giaever G, Flaherty P, Kumm J, Proctor M, Nislow C, Jaramillo DF, Chu AM, Jordan MI, Arkin AP, Davis RW. 2004. Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc Natl Acad Sci USA* 101(3):793–8.

Hancock LC, Behta RP, Lopes JM. 2006. Genomic analysis of the Opi- phenotype. *Genetics* 173(2):621–34.

Hellauer K, Lesage G, Sdicu AM, Turcotte B. 2005. Large-scale analysis of genes that alter sensitivity to the anticancer drug tirapazamine in *Saccharomyces cerevisiae*. *Mol Pharmacol* 68(5):1365–75.

Huang ME, Rio AG, Nicolas A, Kolodner RD. 2003. A genomewide screen in *Saccharomyces cerevisiae* for genes that suppress the accumulation of mutations. *Proc Natl Acad Sci USA* 100(20):11529–34.

Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. *J Comput Graph Stat* 5(3):299–314.

Kawahata M, Masaki K, Fujii T, Iefuji H. 2006. Yeast genes involved in response to lactic acid and acetic acid: acidic conditions caused by the organic acids in *Saccharomyces cerevisiae* cultures induce expression of intracellular metal metabolism genes regulated by Aft1p. *FEMS Yeast Res* 6(6):924–36.

Lam KK, Davey M, Sun B, Roth AF, Davis NG, Conibear E. 2006. Palmitoylation by the DHHC protein Pfa4 regulates the ER exit of Chs3. *J Cell Biol* 174(1):19–25.

Lee W, St Onge RP, Proctor M, Flaherty P, Jordan MI, Arkin AP, Davis RW, Nislow C, Giaever G. 2005. Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS Genet* 1(2):e24.

Lesage G, Sdicu AM, Menard P, Shapiro J, Hussein S, Bussey H. 2004. Analysis of beta-1,3-glucan assembly in *Saccharomyces cerevisiae* using a synthetic interaction network and altered sensitivity to caspofungin. *Genetics* 167(1):35–49.

Lord PW, Stevens RD, Brass A, Goble CA. 2003. Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. *Bioinformatics* 19(10):1275–83.

Marston AL, Tham WH, Shah H, Amon A. 2004. A genome-wide screen identifies genes required for centromeric cohesion. *Science* 303(5662):1367–70.

Mollapour M, Fong D, Balakrishnan K, Harris N, Thompson S, Schuller C, Kuchler K, Piper PW. 2004. Screening the yeast deletant mutant collection for hypersensitivity and hyper-resistance to sorbate, a weak organic acid food preservative. *Yeast* 21(11):927–46.

Ni L, Snyder M. 2001. A genomic study of the bipolar bud site selection pattern in *Saccharomyces cerevisiae*. *Mol Biol Cell* 12(7):2147–70.

Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD. 2004. A robust toolkit for functional profiling of the yeast genome.

Mol Cell 16(3):487–496.

Parsons AB, Lopez A, Givoni IE, Williams DE, Gray CA, Porter J, Chua G, Sopko R, Brost RL, Ho CH, et al. 2006. Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* 126(3):611–25.

Peng WT, Robinson MD, Mnaimneh S, Krogan NJ, Cagney G, Morris Q, Davierwala AP, Grigull J, Yang X, Zhang W, et al. 2003. A panoramic view of yeast noncoding RNA processing. *Cell* 113(7):919–33.

Perrone GG, Grant CM, Dawes IW. 2005. Genetic and environmental factors influencing glutathione homeostasis in *Saccharomyces cerevisiae*. *Mol Biol Cell* 16(1):218–30.

Rong L, Klein HL. 1993. Purification and characterization of the SRS2 DNA helicase of the yeast *Saccharomyces cerevisiae*. *J Biol Chem* 268(2):1252–9.

Reiner S, Micolod D, Zellnig G, Schneiter R. 2006. A genomewide screen reveals a role of mitochondria in anaerobic uptake of sterols in yeast. *Mol Biol Cell* 17(1):90–103.

Resnik P. 1995. Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In: *Proc IJCAI-95*; 1995 Aug 19–20; Montreal. San Mateo (CA): Morgan Kaufmann. p 448–53.

Serrano R, Bernal D, Simon E, Arino J. 2004. Copper and iron are the limiting factors for growth of the yeast *Saccharomyces cerevisiae* in an alkaline environment. *J Biol Chem* 279(19):19698–704.

Shannon CE. 1948. A mathematical theory of communication. *Bell Syst Tech J*

27:379–423, 623–56.

van Voorst F, Houghton-Larsen J, Jonson L, Kielland-Brandt MC, Brandt A. 2006. Genome-wide identification of genes required for growth of *Saccharomyces cerevisiae* under ethanol stress. *Yeast* 23(5):351–9.

Ye P, Peyser BD, Pan X, Boeke JD, Spencer FA, Bader JS. 2005. Gene function prediction from congruent synthetic lethal interactions in yeast. *Mol Syst Biol* 1:2005.0026.

Zewail A, Xie MW, Xing Y, Lin L, Zhang PF, Zou W, Saxe JP, Huang J. 2003. Novel functions of the phosphatidylinositol metabolic pathway discovered by a chemical genomics screen with wortmannin. *Proc Natl Acad Sci USA* 100(6):3345–50.
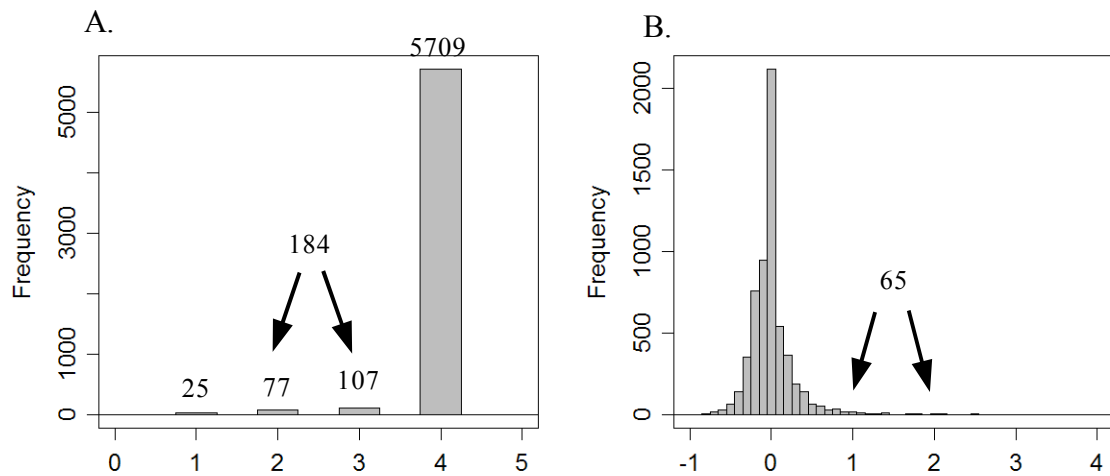
Figure 1. Method for calculating P_i. (A) P_i for gene pair with phenotype values *ORF1* = 2 and *ORF2* = 3 is (77 + 107)/5918. *ORF1–ORF2* phenotype similarity score for this phenotype is $-\log_{10}(184/5918) = 1.5$. (B) P_i for gene pair with *ORF1* = 1.0 and *ORF2* = 2.0 is 65/5918. Phenotype similarity score for this phenotype is $-\log_{10}(65/5918) = 2.0$.

GO Similarity Increases With Phenotypic Similarity
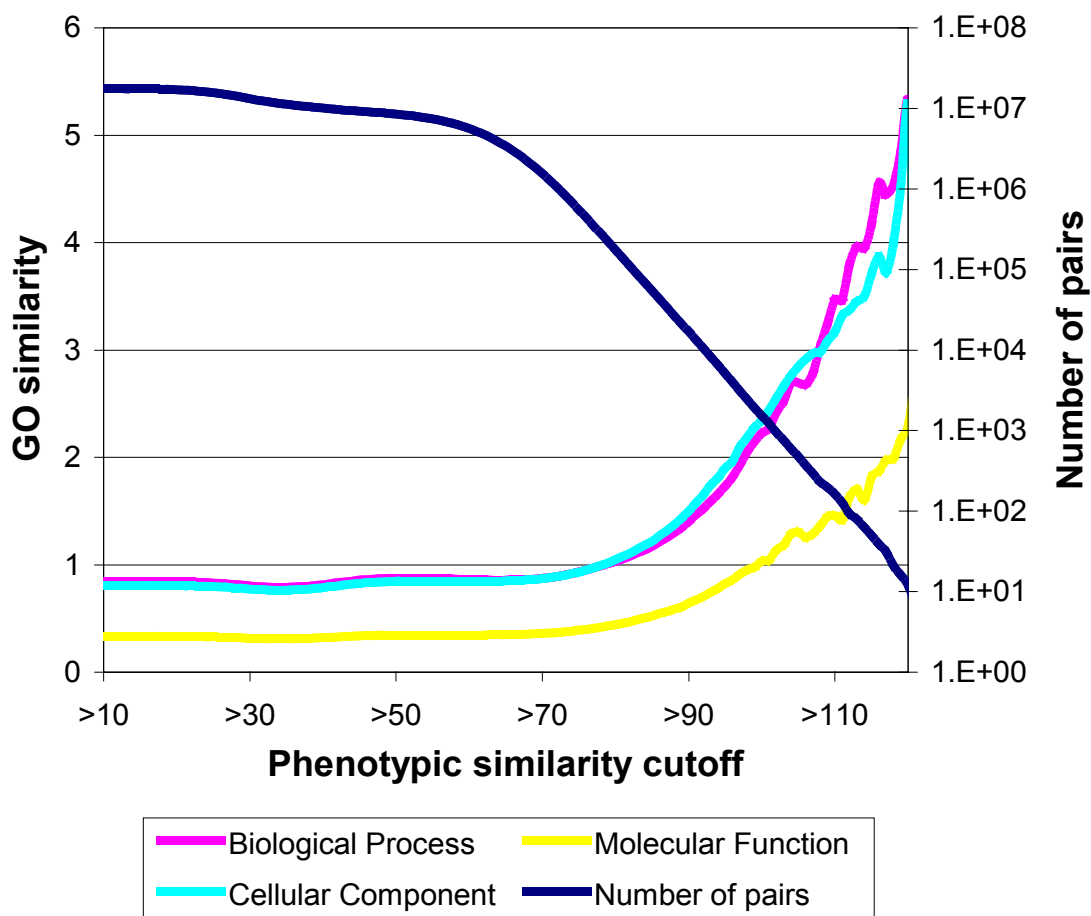
Figure 2. GO similarity increases with phenotype similarity score. The average GO

similarity is shown along increasing phenotype similarity score cutoffs for Biological

Process (magenta line), Cellular Component (cyan line), and Molecular Function (yellow

line). Also shown is the number of gene pairs above each cutoff (blue line).

Enrichment of Interacting Pairs by Phenotypic Similarity
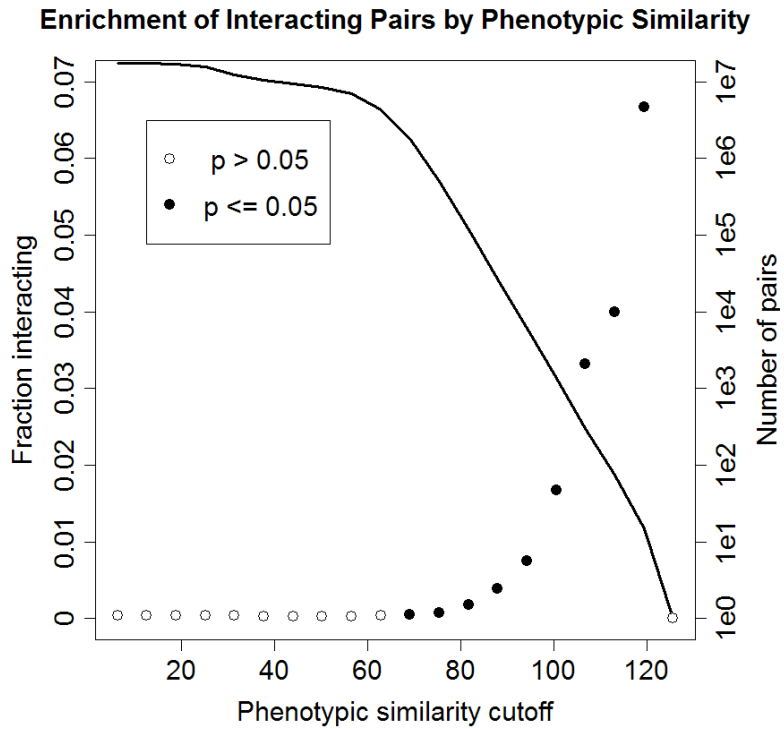
Figure 3. Protein interactions are enriched by phenotypic similarity. The fraction of protein pairs reported to physically interact (Krogan et al. 2006) encoded by gene pairs above various phenotype similarity score cutoffs is shown with circles. Filled circles represent sets that are significantly enriched for interacting pairs ($P \leq 0.05$) by one-sided Fisher's exact test. Line shows number of pairs above each cutoff.

Figure 4. Network generated by connecting all genes with a phenotype similarity score ≥ 100. Dotted line, interaction score 100–105; thin solid line, interaction score 105–110; medium solid line, interaction score 110–115; thick solid line, interaction score > 115. (A) Top half of network. (B) Bottom half of network.
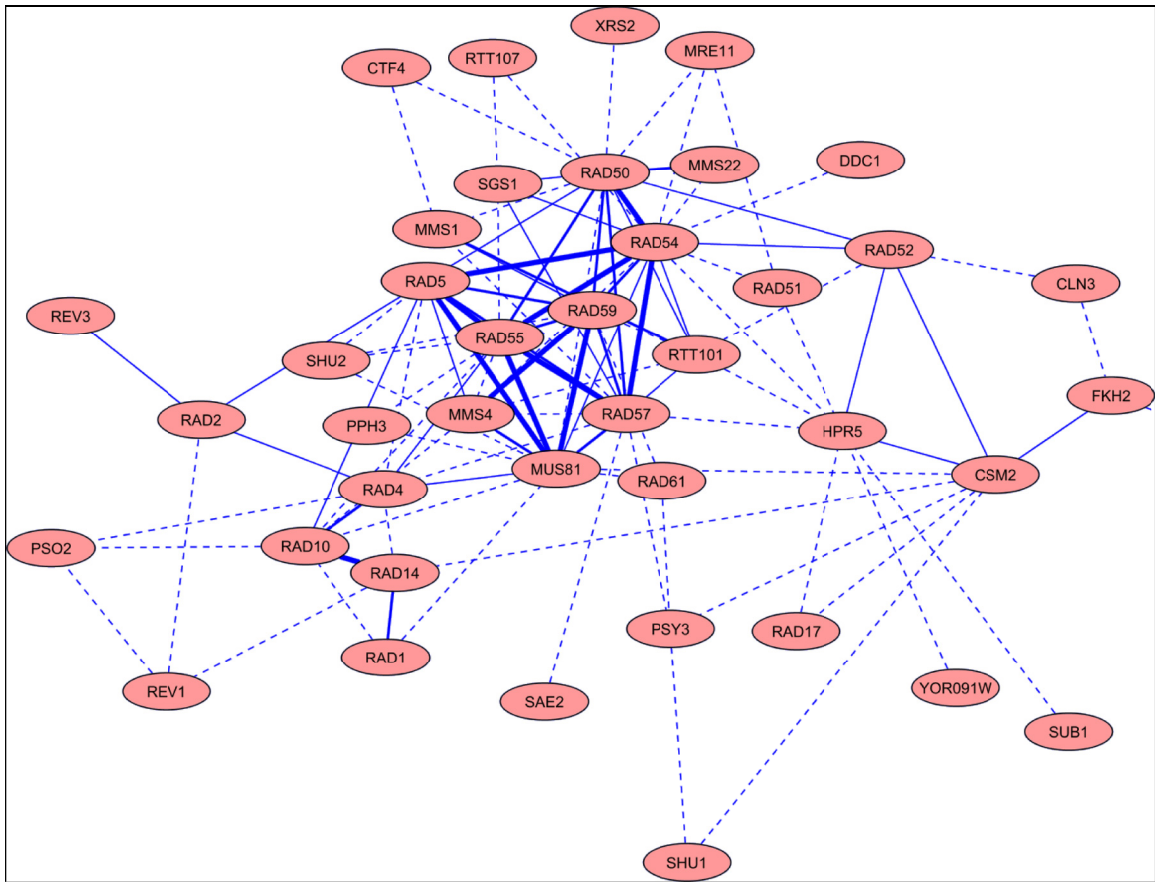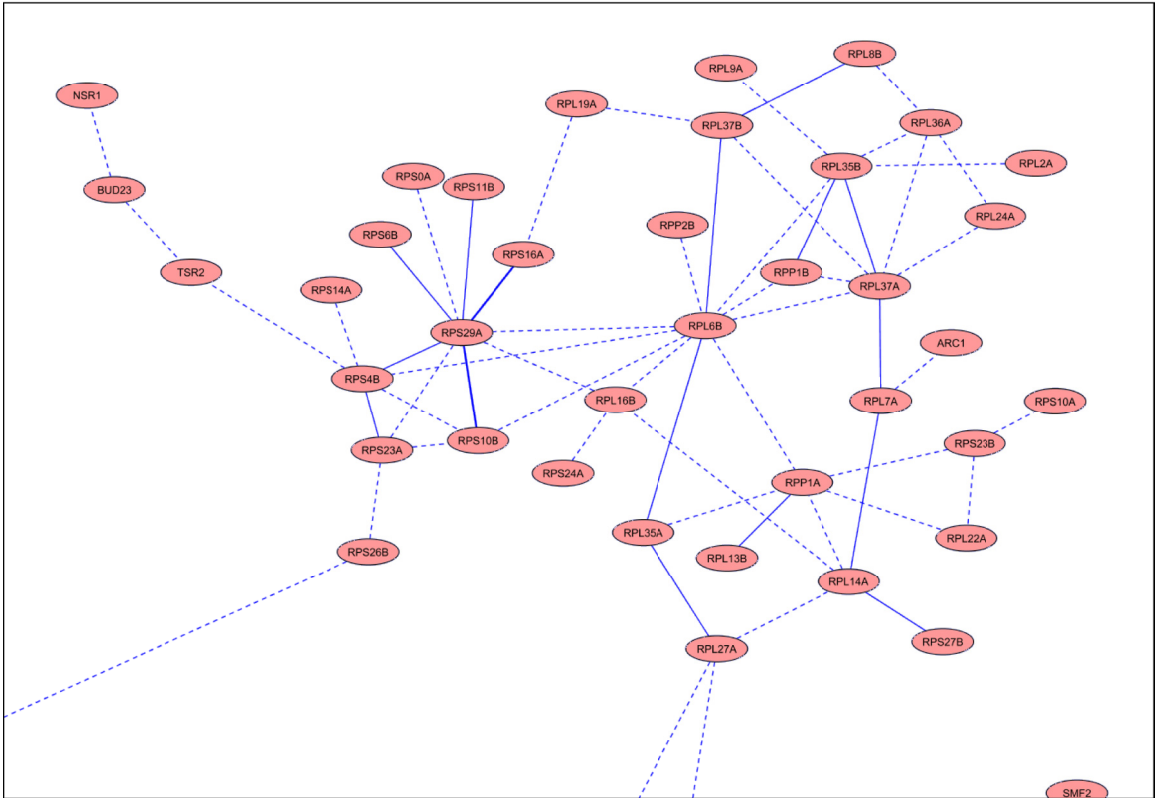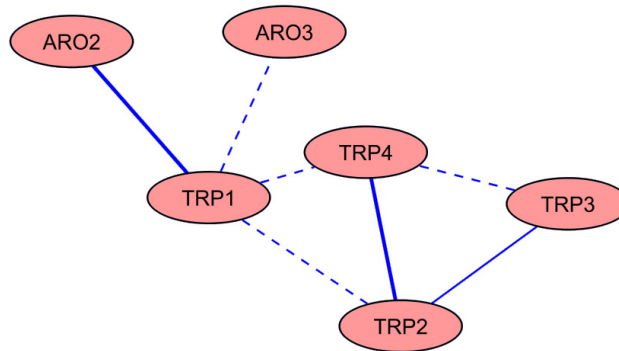
Figure 4A.
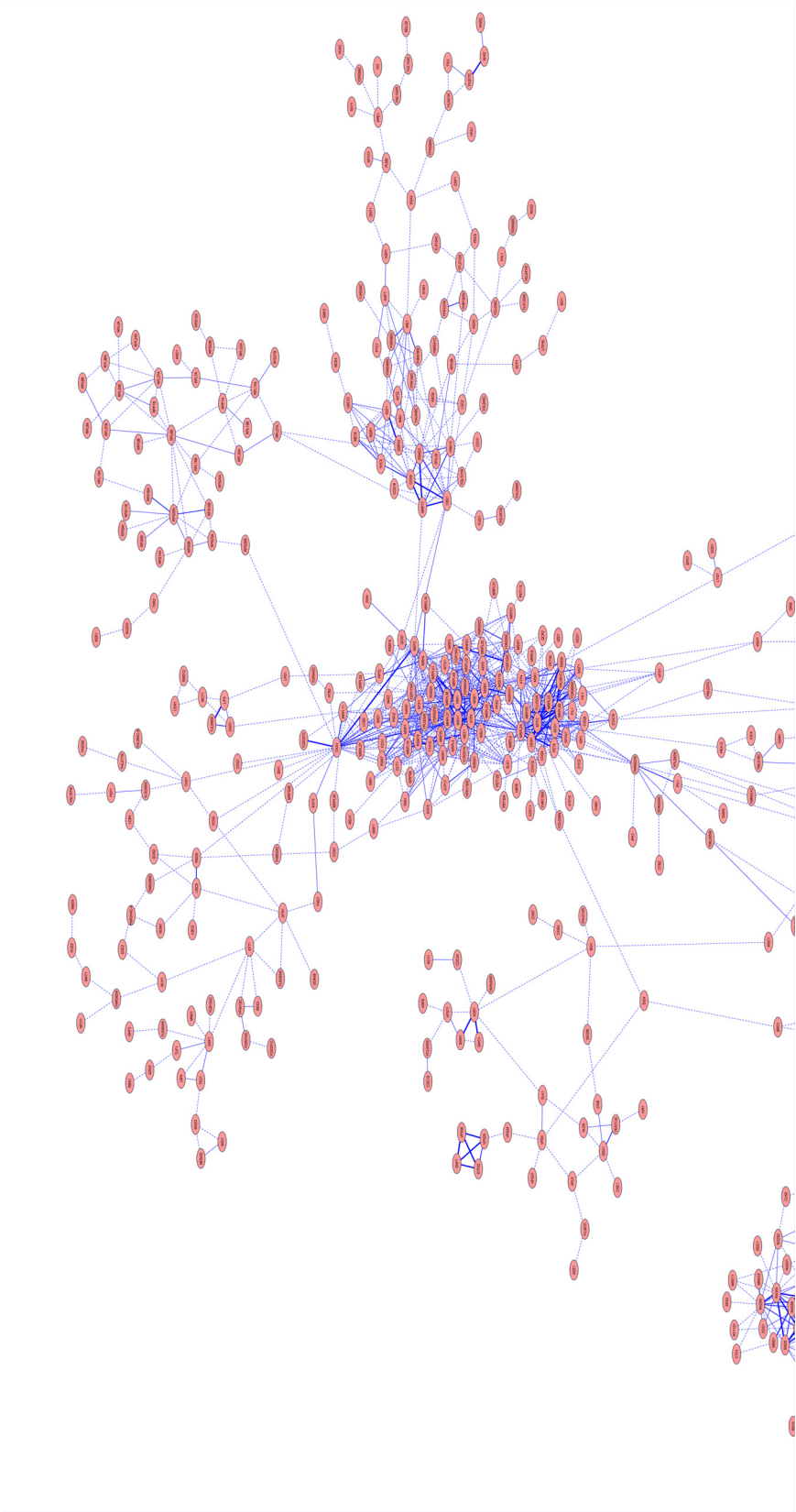
Figure 4B.
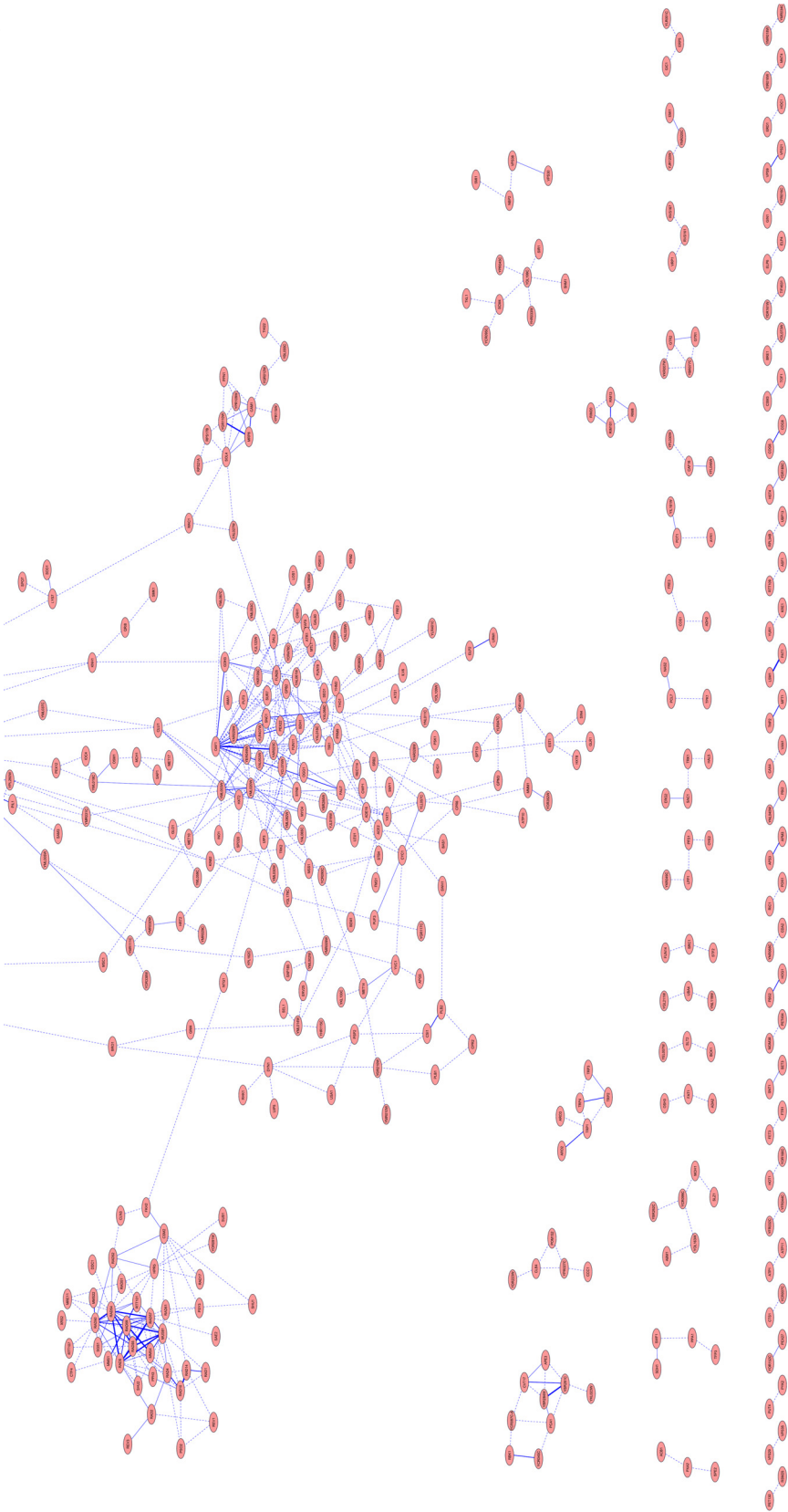
Figure 5. Subset of phenotype similarity score network showing many DNA damage genes. Dotted line, interaction score 100–105; thin solid line, interaction score 105–110; medium solid line, interaction score 110–115; thick solid line, interaction score > 115.

Figure 6. Subset of phenotype similarity score network showing many ribosomal genes. Dotted line, interaction score 100–105; thin solid line, interaction score 105–110; medium solid line, interaction score 110–115.

Figure 7. Subset of phenotype similarity score network showing tryptophan biosynthesis pathway genes. Dotted line, interaction score 100–105; thin solid line, interaction score 105–110; medium solid line, interaction score 110–115.

Table 1. Sources for phenotype data.

Citation	Phenotype
2001-07 Mol Biol Cell Ni L, Snyder M	Homozygous diploid YKOs with unipolar budding pattern
	Homozygous diploid YKOs with axial-like budding pattern
	Homozygous diploid YKOs with random budding pattern: 2=strong, 1=weak
2003-02 Antimicrob Agents and Chemotherapy Blackburn AS, Avery SV	Gentamicin-sensitive YKOs
	oxytetracycline-sensitive YKOs
2003-09-30 Proc Natl Acad Sci USA Huang ME et al.	Rank order of the 33 BY4741 CanR mutator YKOs
2004-02-27 Science Marston AL et al.	% Spores
	% Dyads
	4-spot tetrads
	3-spot tetrads
	2-spot tetrads
2004-05-07 J Biol Chem Serrano R et al.	Alkaline pH sensitive YKOs
2004-05 Genetics Lesage G et al.	Haploid YKOs that show hypersensitivity to caspofungin
	Diploid YKOs that show hypersensitivity to caspofungin
	Haploid YKOs that show enhanced resistance to caspofungin
2004-08 Yeast Mollapour M et al.	YKOs with sensitivity to YPD growth at pH 4.5, relative to pH 6.8
	YKOs with sensitivity to YPD plus 2 mM sorbate pH 4.5
	YKOs with resistance to 5 mM sorbate at pH 4.5
2005-01 Mol Biol Cell Perrone GG et al.	SD medium GSH + GSSG
	SD medium 2× BCAA
	SD medium 4× BCAA
	SD medium pH 6 GSH + GSSG
	SD medium GSSG
2005-06-07 Current Biol Dorer RA et al.	MATa YKOs Most Sensitive to Cincreasin
	MATa/α Heterozygous YKOs Sensitive to Cincreasin
2005-07 Fung Genet Biol Corbacho I et al.	Ldb level
	Invertase Size
2005-11 Mol Pharmacol Hellauer K et al.	YKOs sensitive to 0.2 mM TPZ
	YKOs resistant to 0.5 mM TPZ
2002-07-25 Nature Giaever G et al.	YPG Resistant Log likelihood
	YPG Sensitive Log likelihood
	1.5 M Sorbitol Resistant
	1.5 M Sorbitol Sensitive
	minimal + his/leu/ura Resistant
	minimal + his/leu/ura Sensitive
	1 M NaCl Resistant
	1 M NaCl Sensitive
	pH 8 Resistant
	pH 8 Sensitive
	Nystatin 10 μM Resistant
	Nystatin 10 μM Sensitive
	minimal complete Resistant
	minimal complete Sensitive

Citation	Phenotype
2002-07-25 Nature Giaever G et al.	lys minus Resistant
	lys minus Sensitive
	trp minus Resistant
	trp minus Sensitive
	thr minus Resistant
	thr minus Sensitive
	Lysine Ratio
	Threonine Ratio
	Tryptophan Ratio
	Round
	Size
	Elongate
	Football
	Clumpy
	Chain
	Branch
	Neck abnormalities
2003-03-18 Proc Natl Acad Sci USA Zewail A et al.	Wortmannin sensitivity
2004-01-20 Proc Natl Acad Sci USA Giaever G et al.	5-fluorouracil_19.2 heterozygous YKOs
	5-fluorouracil_38.5 heterozygous YKOs
	5-fluorouracil_4.8 heterozygous YKOs
	5-fluorouracil_76.9 heterozygous YKOs
	5-fluorouracil_9.6 heterozygous YKOs
	alverine-citrate_500 heterozygous YKOs
	atorvastatin_125 heterozygous YKOs
	atorvastatin_62.5 heterozygous YKOs
	cisplatin_125 heterozygous YKOs
	cisplatin_31.25 heterozygous YKOs
	cisplatin_62.5 heterozygous YKOs
	dyclonine_250 heterozygous YKOs
	dyclonine_500 heterozygous YKOs
	fenpropimorph_1.17 heterozygous YKOs
	fenpropimorph_2.34 heterozygous YKOs
	fenpropimorph_4.69 heterozygous YKOs
	fluconazole_130.6 heterozygous YKOs
	fluconazole_16.3 heterozygous YKOs
	fluconazole_32.6 heterozygous YKOs
	fluconazole_65.3 heterozygous YKOs
	lovastatin_154.5 heterozygous YKOs
	lovastatin_77.2 heterozygous YKOs
	methotrexate_125 heterozygous YKOs
	methotrexate_250 heterozygous YKOs
	methotrexate_500 heterozygous YKOs
	miconazole_0.025 heterozygous YKOs
	miconazole_0.05 heterozygous YKOs
	miconazole_0.1 heterozygous YKOs
	miconazole_0.2 heterozygous YKOs
	cisplatin_125 homozygous YKOs

213

Citation	Phenotype
2004-01-20 Proc Natl Acad Sci USA Giaever G et al.	cisplatin_250 homozygous YKOs
	cisplatin_500 homozygous YKOs
	cisplatin_66 homozygous YKOs
2005-08 PLoS Genet Lee W et al.	Cisplatin 500 μM
	Carboplatin 15 mM
	Oxaliplatin 4000 μM
	Psoralen irradiated 0.5 μM
	Angelicin irradiated 62.5 μM
	Mechlorethamine 62.5 μM
	Dmaec 240 000 μM
	Mitomycinc 1 mM
	Streptozotocin 2 mM
	Mms 0.002%
	Camptothecin 30 μg/ml
	Camptothecin 30 000 μg/ml
	4-nqo 0.0313 μM
2006-06 Genetics Hancock LC et al.	Opi- phenotype strains
2006-07-03 J Cell Biol Lam KKY et al.	Low Calcoflour white fluorescence (low chitin)
2006-01 Mol Biol Cell Reiner S et al.	Essential for anaerobic growth but not for aerobic growth
2006-04-15 Yeast van Voorst F et al.	Mutants sensitive to ethanol in unbiased screen
2006-08-04 Cell Parsons AB et al.	-log10(P-val) sensitivity to Sulfometuron methyl
	-log10(P-val) sensitivity to MMS
	-log10(P-val) sensitivity to Clotrimazole
	-log10(P-val) sensitivity to Benomyl
	-log10(P-val) sensitivity to Plumbagin
	-log10(P-val) sensitivity to Hydroxyurea
	-log10(P-val) sensitivity to Artemisinin
	-log10(P-val) sensitivity to Amantadine hydrochloride
	-log10(P-val) sensitivity to 4-Hydroxytamoxifen
	-log10(P-val) sensitivity to Usnic acid
	-log10(P-val) sensitivity to Sodium Azide
	-log10(P-val) sensitivity to Nystatin
	-log10(P-val) sensitivity to Neomycin sulfate
	-log10(P-val) sensitivity to Caffeine
	-log10(P-val) sensitivity to Menthol
	-log10(P-val) sensitivity to Verrucarin
	-log10(P-val) sensitivity to Valinomycin
	-log10(P-val) sensitivity to Trifluoroperazine
	-log10(P-val) sensitivity to Tamoxifen
	-log10(P-val) sensitivity to Raloxifene
	-log10(P-val) sensitivity to Pentamidine
	-log10(P-val) sensitivity to Nigericin
	-log10(P-val) sensitivity to LY-294,002
	-log10(P-val) sensitivity to Latrunculin B
	-log10(P-val) sensitivity to Hydroxyethilhidrazine
	-log10(P-val) sensitivity to Hydrogen peroxide
	-log10(P-val) sensitivity to Hoechst

214

Citation	Phenotype
2006-08-04 Cell Parsons AB et al.	-log10(P-val) sensitivity to Harmine
	-log10(P-val) sensitivity to Haloperidol
	-log10(P-val) sensitivity to Fenpropimorph
	-log10(P-val) sensitivity to Emetine
	-log10(P-val) sensitivity to Dyclonine
	-log10(P-val) sensitivity to Doxycycline
	-log10(P-val) sensitivity to Cyclopiazonic acid
	-log10(P-val) sensitivity to Clomiphene
	-log10(P-val) sensitivity to Cisplatin
	-log10(P-val) sensitivity to Chlorpromazine
	-log10(P-val) sensitivity to Cerulenin
	-log10(P-val) sensitivity to Calcium ionophore
	-log10(P-val) sensitivity to Anisomycin
	-log10(P-val) sensitivity to Amphotericin
	-log10(P-val) sensitivity to Amiodarone
	-log10(P-val) sensitivity to Alamethicin
	-log10(P-val) sensitivity to Actinomycin
	-log10(P-val) sensitivity to Abietic acid
	-log10(P-val) sensitivity to Wortmannin
	-log10(P-val) sensitivity to Staurosporine
	-log10(P-val) sensitivity to Conine
	-log10(P-val) sensitivity to Parthenolide
	-log10(P-val) sensitivity to Radicicol
	-log10(P-val) sensitivity to Mitomycin C
	-log10(P-val) sensitivity to Trichostatin A
	-log10(P-val) sensitivity to FK506
	-log10(P-val) sensitivity to Brefeldin A
	-log10(P-val) sensitivity to U73122
	-log10(P-val) sensitivity to Tunicamycin
	-log10(P-val) sensitivity to Thialysine
	-log10(P-val) sensitivity to Rapamycin
	-log10(P-val) sensitivity to Phenylarsine oxide
	-log10(P-val) sensitivity to Phenantroline
	-log10(P-val) sensitivity to Oligomycin
	-log10(P-val) sensitivity to Nocodazole
	-log10(P-val) sensitivity to Hygromycin B
	-log10(P-val) sensitivity to Extract 95-57
	-log10(P-val) sensitivity to Extract 6592
	-log10(P-val) sensitivity to Extract 00-89
	-log10(P-val) sensitivity to Extract 00-303C
	-log10(P-val) sensitivity to Extract 00-243
	-log10(P-val) sensitivity to Extract 00-192
	-log10(P-val) sensitivity to Extract 00-132
	-log10(P-val) sensitivity to Emodin
	-log10(P-val) sensitivity to Desipramine
	-log10(P-val) sensitivity to Cytochalasin A
	-log10(P-val) sensitivity to CG4-Theopalauamide
	-log10(P-val) sensitivity to Caspofungin
	-log10(P-val) sensitivity to Camptothecin
	-log10(P-val) sensitivity to Basiliskamide
	-log10(P-val) sensitivity to 192A4-Stichloroside
	-log10(P-val) sensitivity to Papuamide B

215

Citation	Phenotype
2006-08-04 Cell Parsons AB et al.	-log10(P-val) sensitivity to Agelasine E
	-log10(P-val) sensitivity to Fluconazole
	-log10(P-val) sensitivity to Geldanamycin
2006-03 FEMS Yeast Res Ando A et al.	Haploid mutants hypersensitive to 30% sucrose
2006-09 FEMS Yeast Res Kawahata M et al.	Haploid mutants resistant to lactic acid
	Haploid mutants resistant to acetic acid
	Haploid mutants resistant to hydrochloric acid
	Haploid mutants sensitive to lactic acid
	Haploid mutants sensitive to acetic acid
	Haploid mutants sensitive to hydrochloric acid

Curriculum Vitæ.

Brian D. Peyser, Ph.D.

March 28, 2007

Personal information:

Born:	Muskingum County, Ohio, USA
	May 21, 1975
Home address:	181 West Main Street
	New Market Maryland 21774
Home telephone:	301-865-3910
Email address:	hopkins@bpeyser.fastmail.net

Education and Experience:

B.S. in Plant Pathology; Microbiology and Cell Science: 1993–1997
        College of Agriculture
        University of Florida
        Gainesville, Florida, USA

Research Assistant: 1993–1997
        Ernest Hiebert, Ph.D., principle investigator
        Department of Plant Pathology
        Institute of Food and Agricultural Sciences
        University of Florida
        Gainesville, Florida, USA

Predoctoral Fellow: 1997–1998
        Cell and Molecular Biology
        School of Medicine
        University of Pennsylvania
        Philadelphia, Pennsylvania, USA

Senior Laboratory Technician: 1998–2000
Biological Scientist: 2000–2002
        Maureen M. Goodenow, Ph.D., principle investigator
        Department of Pathology and Laboratory Medicine
        College of Medicine
        University of Florida
        Gainesville, Florida, USA

Ph.D. in Biochemistry, Cellular and Molecular Biology: 2002–2007
    Forrest A. Spencer, Ph.D., advisor
    Department of Molecular Biology and Genetics
    School of Medicine
    The Johns Hopkins University
    Baltimore, Maryland, USA

Publications:

1. **Peyser, B. D.**, P. B. Meluh, C. W. Tiffany, O. Chen, D. S. Yuan, S. Sookhai-Mahadeo, E. Kleinhans, A. Khan, J. D. Boeke, F. A. Spencer, and R. A. Irizarry. 2007 in preparation. Synthetic lethal interactions predicted by analysis of high-throughput dSLAM.

2. **Peyser, B. D.**, R. Irizarry, and F. A. Spencer. 2007 in press. Statistical analysis of fitness data determined by TAG hybridization on microarrays. In: *Gene Essentiality at Genome Scale: Protocols and Bioinformatics* (A. Osterman and S. Gerdes, Eds.), Totowa, NJ.

3. Ooi, S. L., X. Pan, **B. D. Peyser**, P. Ye, P. B. Meluh, D. S. Yuan, R. A. Irizarry, J. S. Bader, F. A. Spencer, and J. D. Boeke. 2006. Global synthetic-lethality analysis and yeast functional profiling. *Trends in Genetics* 22(1):56–63.

4. **Peyser, B. D.** and P. Ye, X. Pan, J. D. Boeke, F. A. Spencer, and J. S. Bader. 2005. Gene function prediction from congruent synthetic lethal interactions in yeast. *Molecular Systems Biology* 1(November 22):msb4100034-E1–12.

5. Ye, P., **B. D. Peyser**, F. A. Spencer, and J. S. Bader. 2005. Commensurate distances and similar motifs in genetic congruence and protein interaction networks in yeast. *BMC Bioinformatics* 6(270).

6. **Peyser, B. D.**, R. A. Irizarry, C. W. Tiffany, O. Chen, D. S. Yuan, J. D. Boeke, and F. A. Spencer. 2005. Improved statistical analysis of budding yeast TAG microarrays revealed by defined spike-in pools. *Nucleic Acids Research* 33(16):e140.

7. Yuan, D. S., X. Pan, S. L. Ooi, **B. D. Peyser**, F. A. Spencer, R. A. Irizarry, and J. D. Boeke. 2005. Improved microarray methods for profiling the yeast knockout strain collection. *Nucleic Acids Research* 33(12):e103.

8. Warren, C. D., D. M. Eckley, M. S. Lee, J. S. Hanna, A. Hughes, **B. Peyser**, C. Jie, R. Irizzary, and F. A. Spencer. 2004. S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion. *Molecular Biology of the Cell* 15(4):1724–35.

9. Abouzid, A. M., J. Freitas-Astua, D. E. Purcifull, J. E. Polston, K. A. Beckham, W. E. Crawford, M. A. Petersen, **B. Peyser**, C. Patte, and E. Hiebert. 2002. Serological Studies Using Polyclonal Antisera Prepared Against the Viral Coat Protein of Four Begomoviruses Expressed in *Escherichia coli. Plant Disease* 86(10):1109–14.

10. Perez, E. E., S. L. Rose, **B. D. Peyser**, S. L. Lamers, B. R. Burkhardt, B. M. Dunn, A. D. Hutson, J. W. Sleasman, and M. M. Goodenow. 2001. Human immunodeficiency virus type 1 protease genotype predicts immune and viral responses to combination therapy with protease inhibitors (PIs) in PI-naive patients. *Journal of Infectious Diseases* 183(4):579–88.