Quantitative Yeast Phenomics: Systematic Single Cell Analysis of DNA Damage Foci

Erin B. Styles [1], Karen J. Founk [1,2], Lee A. Zamparo [1,2,3], Tina L. Sing [1,4], Dogus Altintas [5], Cyril Ribeyre [5,6], Virginie Ribaud [5,7], Jacques Rougemont [8], David Mayhew [9], Michael Costanzo [1], Matej Usaj [1], Adrian J. Verster [1], Elizabeth N. Koch [10], Daniele Novarina [11], Marco Graf [12], Brian Luke [12], Marco Muzi-Falconi [11], Chad L. Myers [10], Robi David Mitra [9,] David Shore [5], Grant W. Brown [1,4], Zhaolei Zhang [1,2], Charles Boone [1,2], Brenda J. Andrews [1,2]

[1] The Donnelly Centre, [2] Department of Molecular Genetics, University of Toronto, Toronto, [3] Department of Computer Sciences, University of Toronto, Toronto, [4] Department of Biochemistry, University of Toronto, Toronto, Ontario Canada M5S 3E1

[5] Department of Molecular Biology, NCCR Program "Frontiers in Genetics", and Institute of Genetics, Genomics, Geneva (iGE3), University of Geneva, 30, quai Ernest-Ansermet, CH-1211, Geneva 4, Switzerland

[6] Current address: Institute of Human Genetics, CNRS UPR 1142, 141 rue de la Cardonille, 34396 Montpellier Cedex 5 France

[7] Current address: Centre Medicale Universitaire, University of Geneva, rue Michel-Servet 1, CH-1211 Geneva 4, Switzerland

[8] Laboratory of Computational Systems Biology, Ecole Polytéchnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[9] Department of Genetics and Center for Genome Sciences and Systems Biology, Washington University School of Medicine in St. Louis, St. Louis, MO 63108

[10] Department of Computer Science and Engineering, University of Minnesota, Minneapolis, Minnesota, USA

[11] Dipartimento di Bioscienze Universita' degli Studi di Milano, Milano, Italy

[12] Institute of Molecular Biology (IMB), Ackermannweg 4, Mainz 55128, Germany

Correspondence to: Charles Boone or Brenda J Andrews
charlie.boone@utoronto.ca, brenda.andrews@utoronto.ca

Summary

We developed an integrated method that combines genome-wide genetic perturbation with high-content screening to enable a complete genetic description of sub-cellular structures and compartment morphology. We exploit automated yeast genetics to create arrays of mutant strains that express fluorescent markers of interest for automated image analysis. As proof-of-principle, we used a Rad52-GFP marker to examine DNA damage foci in ~20 million single cells from ~5000 different mutant backgrounds in the context of selected genetic or chemical perturbations. We used pattern classification through machine learning to identify 345 mutants that had elevated levels of DNA damage foci, almost half of which were identified only in sensitized backgrounds. We discovered a role for the BED-domain-containing protein Vid22 and the Sgs1 DNA helicase at sites of DNA damage, and for Vid22 at predicted G-quadruplex regions. Our approach is extensible to numerous other cell biological markers and experimental systems.

Introduction

A fundamental goal of functional genomics projects is to systematically define cellular pathways and networks. In the budding yeast *Saccharomyces cerevisiae*, genome-wide collections of haploid viable deletion mutants (Giaever et al., 2002; Winzeler et al., 1999) and mutant strains carrying conditional alleles of essential genes (Kofoed et al., 2015; Li et al., 2011; Mnaimneh et al., 2004) enable systematic genetic analysis. Due to ease of measurement and amenability to high-throughput applications, most genome-scale studies have focused on cell fitness as a phenotypic readout (Baryshnikova et al., 2010; Collins et al., 2007). Notably, colony size has been used to examine the fitness phenotype of millions of double mutant gene pairs to produce a global yeast genetic interaction network (Costanzo et al., 2010). Despite the information-rich nature of fitness assays, it is clear that the analysis of more subtle and specific phenotypes will yield important new functional information. For example, while ~10% of the non-essential yeast deletion mutants show a clear fitness defect, nearly ~50% exhibit a number of different morphological defects (Ohya et al., 2005). Thus, comprehensive understanding of gene function and genetic interaction networks will require further analysis of more complex phenotypes in a variety of conditions to reveal a complete functional wiring diagram of the cell.

In the past decade, systematic assessment of subcellular spatio-temporal phenotypes using high-content screening approaches has emerged as a powerful approach for functional analysis (Carpenter et al., 2006). While several studies have examined

subcellular morphology systematically in yeast (Alvaro et al., 2007; Breker et al., 2013; Chong et al., 2015; Huh et al., 2003; Tkach et al., 2012), a global characterization of mutant phenotypes remains a major challenge. Here we describe a high-throughput pipeline for quantifying mutant phenotypes by combining two automated platforms: synthetic genetic array (SGA) analysis, which automates yeast genetics (Tong et al., 2001), and high-content screening (HCS), which enables quantitative cell biological analysis at the single cell level (Chong et al., 2015; Li et al., 2011; Vizeacoumar et al., 2010). Our method provides a fully scalable approach for systematic analysis of yeast cell images, enabling the detection of subcellular morphological defects in response to thousands of genetic perturbations, in a quantitative and statistically robust manner.

As a case study, we monitored the presence of a transient, gigadalton-sized assembly of proteins referred to as a DNA damage-induced focus. This subnuclear complex arises in response to double stranded DNA breaks, and acts as a recombination center for DNA repair (Lisby et al., 2001). The DNA damage focus is an appealing compartment for development of an automated imaging pipeline for several reasons. First, key proteins that form and influence the focus are highly conserved and many genes with potential roles in focus formation or regulation have been identified through manual image inspection, providing useful positive controls (Alvaro et al., 2007). Second, the focus is a relatively simple shape, is usually found as a single entity in the cell, and has a substantial half life (approximately five minutes; Lisby et al., 2003), facilitating the development of useful statistical approaches and automated imaging protocols. We

used the DNA damage focus marker, Rad52-GFP (Alvaro et al., 2007), to score foci formation in thousands of non-essential gene deletion mutants (Giaever et al., 2002; Winzeler et al., 1999) and conditional temperature-sensitive (TS) alleles of essential genes (Li et al., 2011), both in the presence and the absence of environmental and genetic perturbations. Our general approach is readily adaptable to other cell biological markers and experimental systems, and enables systematic and quantitative analysis of genes influencing subcellular compartment morphology or pathway activity.

Results

Designing a robust SGA-High-Content Screening pipeline for identification of cell populations with elevated levels of DNA damage foci

Our strategy for systematic phenotypic analysis of the DNA damage focus required development of a completely automated pipeline for imaging and scoring DNA damage foci phenotypes in thousands of different yeast mutants. The first component of the pipeline involved assembly of yeast mutant arrays compatible with high-throughput (HTP) image acquisition and analysis. We constructed SGA-compatible yeast strains containing phenotypically neutral markers for fluorescently labeled DNA damage foci (*RAD52-GFP*; Figures 1A and S1A,B), nuclei (*HTA2-mCherry*; Figures 1A and S1B) and the cytoplasm (*RPL39*pr-*tdTomato;* Figure 1A). SGA and high-content screening tools were then used to visualize fluorescent proteins in the yeast nonessential gene deletion collection, as well as a collection of mutants carrying TS alleles of essential genes (Figure 1A,B; Li et al., 2011). Three yeast mutant arrays were constructed using this protocol: [1]

a single mutant array which was assessed in the presence and absence of the DNA damaging agent phleomycin; [2] a double mutant array in which each strain carried a deletion allele of *SGS1*, which encodes a nonessential DNA helicase; [3] a double mutant array deleted for *YKU80*, which encodes a nonessential protein involved in non-homologous end-joining and telomere maintenance. We reasoned that mutation of *SGS1* or *YKU80* would sensitize the cell to defects in the DNA damage response in different ways, thus expanding our ability to discover a diverse set of functionally relevant genes.

The second component of our SGA-HCS pipeline involved HTP microscopy, and automated image analysis and pattern classification through machine learning. Cell boundaries and nuclei were identified in the red channel, and 470 features were extracted from both the red and green channels for each cell using CellProfiler (Carpenter et al., 2006). We then coupled our feature selection with support vector machine-based (SVM) machine learning to generate a classifier capable of distinguishing nuclei with at least 1 focus from nuclei lacking any foci (Supplemental Experimental Procedures; Figure S1C-E).

The final component of our method exploits the systematic and automated nature of the screening pipeline to address critical statistical considerations that could not be addressed in previous studies. Specifically, we defined three parameters to determine a cutoff at which biologically relevant hits could be identified: [1] a minimal cell sample

size for reliable measurement; [2] a score to identify mutants that differ significantly from wild-type cells; [3] a normalization strategy to remove screening bias. First, to determine the sample size required for reliable scoring of yeast mutants for foci detection, staggered sizes of populations were randomly selected from a pool of ~170,000 Rad52-GFP cells and scored for the presence of foci (Figures 1D and S1F). This analysis revealed that an imaging sample size of 1000 cells/mutant allowed for reliable measurements of foci frequency (Standard deviation = 0.82%), a sample size that is difficult to achieve using manual assessment. The sample size criterion was met for ~80% of mutants examined in our experimental pipeline (Figure 1E); mutants with a severe fitness defect often failed this step and strains for which fewer than 250 cells were observed across all biological replicates were excluded from further analysis.

We developed a score to reliably identify mutants that accounts for batch effects and other experimental biases that are typical of large-scale screens. We first scored and ranked mutants using a binomial distribution, a statistical test that determines the likelihood of a mutant having the same fraction of cells with foci as wild-type in the context of sample size, and replicates were combined to generate a single score for each mutant using Fisher's method (Elston, 1991; Skellam, 1948). To eliminate plate-specific effects, scores were normalized to the average fraction of foci on each plate, rather than a global average. The B-score, a non-parametric measure of deviation analogous to the well-known Z-score, allowed us to filter out hits that contained bias as a result of positional effects within single plates (Malo et al., 2006). A combination of the binomial

test and B-score out-performed the binomial test alone based on functional enrichment analysis of rank-ordered single mutants (Figure 1F; sample size interval = 10; Max sample set = 200). Top-ranking single mutants scored using this method were more enriched for genes involved in DNA repair, homologous recombination, and cohesion than single mutants scored using raw foci percentage or Z-score.

Applying SGA-HCS to map networks of DNA damage response genes

We used our optimized SGA-HCS pipeline to observe ~1000 yeast cells in each of ~5000 different mutant backgrounds, in the context of four separate genetic and/or chemical perturbations for a total of ~24,000 different mutant populations and ~20 million single cells. On average, ~7% of the individual cells within a wild-type population exhibited a single focus, likely due to stalled replication forks and other endogenous sources of DNA damage (Figure 1C,D); however, application of our scoring criteria identified 345 loss-of-function mutants that had elevated levels of DNA damage foci, either as single mutants or as double mutants, when combined with the deletion of *SGS1*, *YKU80*, or following treatment with a DNA damaging agent, phleomycin (Figure 2; Table S2). Almost half of the mutants (48%) were identified only in a sensitized background or condition, consistent with previous work illustrating the importance of considering genetic or chemical-genetic interactions for optimal exploration of yeast pathways (Costanzo et al., 2010; Vizeacoumar et al., 2010). In fact, each chemical or genetic sensitization experiment identified a distinct set of mutants with elevated levels of DNA damage foci. For example, mutants identified in the absence of Yku80 were uniquely enriched for

those with abnormal telomere size (*P-value* = 9.05×10^{-07}; Figure 3A; Askree et al., 2004), consistent with known functions for Yku80. In contrast, enzymes involved in DNA metabolism, including the exonuclease Exo1, the endonuclease Mus81, and the helicase Srs2, were uniquely detected in the *sgs1*Δ double mutant screen (Figure 2), which contained numerous genes that show an *SGS1* genetic interaction, or whose products are known to interact physically with Sgs1 (32/44; Figure S2A). A subset of these hits (14/44 non-essential mutants) showed increased sensitivity to HU or MMS in the *sgs1*Δ background, supporting a combined role with Sgs1 in the DNA damage response (DDR; Figure S2B,C). Importantly, our phenotypic analysis provides new information that would be missed in fitness-based genetic interaction studies because 51% of the hits identified in the *sgs1*Δ screens were not previously reported to show a genetic interaction, despite extensive SGA genetic analysis with an *sgs1*Δ query strain (Figure 2; Costanzo et al., 2010).

We validated the results of our primary screens with two secondary assays, one using an independent deletion mutant collection to score Rad52 foci and another involving a test of plasmid-based gene complementation (Ho et al., 2009). Although we were not able to test all primary hits in the secondary assays due to technical issues (e.g. if a specific plasmid or strain was not available), we confirmed 152 mutants of 230 tested in at least one validation assay, which is suggestive of an upper boundary for the false positive rate of ~30% (Figure S3A-E; Costanzo et al., 2010). False positives may reflect discrepancies between biological replicates or aberrant cell segmentation issues with the original

screen (Figure S1). We estimated a similar false negative rate of ~30% by assessing discrepancies amongst Rad52 focus phenotype in strains mutated for genes encoding members of the same protein complex, which should behave similarly in general (Figure S3F; Table S3).

We used several comparative analyses to validate our primary screen results. First, we identified many genes with known roles in the DDR in the primary hit list from our screens. For example, genes important for double strand break (DSB) repair were detected, including all tested members of the Rad52 epistasis group, DNA replication genes, and genes important for activating the DNA damage checkpoint (Figure 2). Second, genes identified in our screens showed enrichment for DNA replication, DNA repair, homologous recombination and cohesion functions (Figure 3A and Table S4). Third, hits from our non-essential mutant screens showed significant overlap with published screens that assayed single deletion mutants for sensitivity to DNA-damaging agents, higher frequencies of chromosomal loss, changes in telomere length, and increased DNA mutation rates (Figure 3A and Table S4). Similarly, non-essential deletion mutants identified in our screens were enriched for genes with genetic interaction profiles resembling those annotated to functional categories affiliated with the DDR (Costanzo et al., 2010), indicating that these genes also have a DDR role (Figure 3B). Finally, we performed a direct comparison between the single deletion mutants identified in this study and those found in a previous manual screen for increased Rad52-YFP foci (Alvaro et al., 2007); our study detected 31% of previously identified

genes as well as 101 unique hits (Figure S4A). Importantly the genes uniquely identified in our study are enriched for DDR-associated functions (LOD = 1.04, *P-value* = 4.38×10^{-16}), indicating that this pool is likely enriched for true positives, which may have been missed due to sample size issues and biases associated with manual image assessment.

The number of foci observed in populations of non-essential mutants showed statistically significant associations with several physiological and evolutionary properties of yeast genes. Notably, we identified a strong correlation between an elevated foci phenotype and single mutant fitness, wherein mutants with increasingly severe foci phenotypes tend to have severe fitness defects (Figure S4B). We also observed correlations between a number of gene and protein attributes of biological networks (e.g. number of protein-protein and genetic interactions, evolutionary conservation, phenotypic capacitance) and the elevated foci phenotype, consistent with previous observations that genes involved in DNA maintenance and organization have broad phenotypic impact (Figure S4B; Levy and Siegal, 2008).

Roles for Vid22 in Promoter Binding, DNA damage response, and G-quadruplex DNA

To explore novel DDR biology in our network, we decided to focus on Vid22, whose deletion caused a dramatic elevated focus phenotype only in the absence of Sgs1. Vid22 was recently linked to the DDR (Bonetti et al., 2012), but a potential functional relationship between Vid22 and Sgs1 was largely unexplored. Vid22 contains a BED

(BEAF and DREF; boundary element-associated factor and DNA replication-regulated element binding factor, respectively)-finger domain, consistent with a function involving DNA-binding (Aravind, 2000). Vid22 physically interacts with both Tbf1, an essential Myb domain telomere binding protein, and Env11, another BED-finger domain protein and paralog of Vid22 (Preti et al., 2010). These three proteins form a stable complex in which Vid22 and Env11 serve to stabilize the chromatin association of Tbf1 (Preti et al., 2010; Ribaud et al., 2012), and promote nucleosome rearrangements around promoters (Badis et al., 2008; Preti et al., 2010) and DSBs (Bonetti et al., 2012).

To investigate the biochemical connection between Sgs1 and Vid22, we used chromatin immunoprecipitation (ChIP) to assay recruitment of a Myc-tagged version of Vid22 to a unique, induced DNA double strand break (Figure 4A,B; Ribeyre and Shore, 2012). Vid22-Myc was strongly recruited (~60-fold enriched over background) to both sides of an induced DSB, consistent with previous work (Bonetti et al., 2012). Notably, recruitment of Vid22 was entirely dependent on *SGS1* (Figure 4C). Because we also observed a genetic interaction between *VID22* and *SGS1*, these two proteins may work in concert to control critical DNA repair functions during normal cell growth. We next performed a kinetic analysis to examine focus formation evident in *vid22*Δ *sgs1*Δ mutants, tracking cells over the course of eight hours (Figure 4D). While the wild-type and *sgs1*Δ cells formed either one or two large foci, both *vid22*Δ and *vid22*Δ *sgs1*Δ cells formed multiple smaller foci in each cell (Figure 4E and Movie S1A-D), with the phenotype being more pronounced in the double mutant. We also observed a higher frequency of long lasting, unresolved foci in these mutant strains (foci lasting longer

than three hours; Figure 4F), consistent with a pronounced defect in DNA damage repair (Lukas et al., 2011).

We next explored aspects of genome integrity and DNA repair mechanisms in the mutant cells. First, we assayed the effect of *VID22* and/or *SGS1* deletion on the integrity of the rDNA cassette, a series of 9.2kb repeat sequences of Chromosome XII that are sequestered into the nucleolus; Rad52 is normally excluded from the nucleolus to prevent recombination between the repeats (Torres-Rosell et al., 2007). We employed an *ADE2* reporter system to assay unequal sister chromatid exchange (USCE) within the rDNA cassette, and saw an increase in marker loss in *vid22*Δ *sgs1*Δ cells, beyond that observed in wild-type or single mutant strains (Figure S5A,B). Furthermore, we saw elevated levels of extrachromosomal circles (ERCs) caused by intra-chromosomal recombination involving the rDNA in *vid22*Δ, *sgs1*Δ, and *vid22*Δ *sgs1*Δ double mutant strains (Figure S5C,D). Consistent with these phenotypes, sub-nucleolar Rad52-GFP foci were increased in *vid22*Δ *sgs1*Δ cells (Figure S5E), which may also explain the elevated levels of DNA damage-induced foci identified in our primary screens. We also assessed DNA damage phenotypes in *vid22*Δ mutants using a series of strains featuring galactose-inducible HO breaks in different genetic contexts, each of which query a different aspect of the DNA repair pathway. Growth of mutant strains was assessed using a serial spot dilution, and revealed that *vid22*Δ mutants exhibited decreased fitness in a strain featuring a non-homologous end joining (NHEJ) repair-dependent HO break (Figure S6), consistent with previous results linking Vid22 to this DNA repair mechanism (Bonetti et

al., 2012).

As noted above, Vid22 and its paralog Env11 localize to promoter sites throughout the genome, together with the general regulatory factor Tbf1 (Preti et al., 2010). We wondered about the relationship between the Sgs1-dependent role for Vid22 in the DDR and the function(s) for Vid22 at promoter regions. We also explored the sites of Vid22 binding throughout the genome using a calling card assay, which assesses the frequency and location of Vid22-mediated Ty5 transposon integration on chromosomal DNA *in vivo* (Wang et al., 2011). Consistent with our ChIP-Seq analysis (Figure 5A; 67.5 fold increase over expected, *P-value* = 2.9^{-56}; Preti et al., 2010), Vid22 localized specifically in promoter regions of 161 genes (Table S5), including those involved in DNA replication, repair and recombination (e.g. *RAD14, NSE1, HUG1*), and genes involved in ER-Golgi trafficking (e.g. *SED4, FRT2*), consistent with work suggesting a secretory defect in *vid22*Δ mutants (Brown et al., 2001; 2002). Although the Tbf1, Vid22 and Env11 ChIP-Seq sites at non-snoRNA promoters show high overlap (Figure 5B and Table S5; Preti et al., 2010), Vid22 recruitment to representative promoters was not affected by *SGS1* deletion (Figure 5C). Thus, the role of Vid22 and Sgs1 in the DDR appears distinct from the function of Vid22 in gene regulation.

Detailed analysis of our global Vid22 calling card and ChIP-Seq analyses revealed an intriguing enrichment for predicted G-quadruplex (G4) DNA regions at Vid22 binding sites (Figure 5D; 6.9 fold increase over expected, *P-value* = $1.6\text{x}10^{-19}$, and 5.2 fold

increase over expected, *P-value* = $1.7x10^{-13}$, respectively; Capra et al., 2010). G4 DNA regions, have the potential to form four-stranded G4 quadruplex structures which are predicted to result from the opening of the DNA helix during either replication or transcription, and are resolved by a number of helicases, including Sgs1 (Reviewed in Maizels and Gray, 2013). G4 DNA that is formed on the non-template strand during transcription is associated with a stable co-transcriptional RNA-DNA hybrid on the template strand, and is highly susceptible to DNA damage. Interestingly, the Vid22 ChIP and calling card binding sites were slightly enriched at loci known to be susceptible to RNA-DNA hybrids in an *rnh1*Δ *rnh201*Δ background (Figure 5D; 1.5 fold increase, *P-value* = 0.02, and 1.4 fold increase, *P-value* = 0.05, respectively; Chan et al., 2014)**.** This suggests a possible functional overlap between Vid22 and members of the RNase HI and/or RNase HII complexes, which remove RNA-DNA hybrids by degrading RNA (reviewed in Aguilera and García-Muse, 2012), in addressing these hybrids.

Since Sgs1 can function as a G4 helicase (Sun et al., 1999), it is possible that Vid22 somehow facilitates the unwinding of G4 structures, or assists in the removal or prevention of the stable RNA-DNA heteroduplex at G4 sites. To investigate these possible functions, we examined genetic interaction data to ask which mutant strains share genetic interactions in common with a *vid22*Δ strain, a phenotype that is typical of genes that function in similar biological processes and pathways (Costanzo et al., 2010). As expected, *VID22* shared many genetic interactions with genes involved in the DDR, including negative genetic interactions with the structural maintenance of

chromosomes (SMC) complex (Figure 6A), which has key roles in DNA repair and the segregation of repetitive DNA regions (Torres-Rosell et al., 2005). Consistent with a possible relationship between Vid22 and RNA-DNA hybrids, the same set of genetic interactions was also seen in strains mutated for genes involved in removal of RNA primers from DNA, including *RNH201*, *RNH202*, and *RNH203,* which encode the members of the RNase HII complex, as well as *DNA2,* which encodes a helicase and tracking protein for flap cleavage during Okazaki fragment maturation (Figure 6A). Interestingly, the apparent functional relationship between RNase HII and Vid22 that is suggested by these genetic interaction profiles was recapitulated in our SGA-HCS analysis because we observed that RNase HII is required for genome integrity in an *sgs1*Δ mutant background. Consistent with the genetic data, we observed localization of RNase HII to induced DSBs (Figure 6B), and RNHII mutant strains had increased Rad52 foci in an *sgs1*Δ mutant background (Figure 2). These results are suggestive of a possible shared but complementary role for Vid22 and the RNase HII complex at sites of DNA damage.

Discussion

In this study, we describe development of an optimized pipeline for combining SGA analysis with high content screening to identify budding yeast mutants with aberrant subcellular morphology, using the DNA damage response pathway as a case-study. We focused on developing methodology for identifying significant mutant phenotypes in cell images. As for other functional genomics screens (Baryshnikova et al., 2010),

normalization of batch effects, including plate-specific results and spatial effects within a microtitre plate that influence the fraction of DNA damage-induced foci in a given population, was critical for a statistically robust measurement of the focus phenotype of each mutant. SGA and liquid handling for HCS are automated separately in our pipeline, making these experimental factors somewhat sporadic, and not amenable to classical normalization approaches which treat all data with a single correction factor (Malo et al., 2006). Filtering data based on a site-specific correction factor that takes into account the relative spatial effect incurred at each location within the context of all other mutants on each plate was key for distinguishing biological effects as opposed to experimental anomalies. Our normalization protocol requires consolidation of data from multiple biological replicates, and the identification of a minimum number of cells to be counted for statistical reliability, which is easily achieved with computational image analysis.

We chose the DDR for optimization of our integrated SGA-HCS pipeline, since the core biology of DNA repair is well-studied and conserved (Lisby et al., 2004), yet recent efforts to explore the DDR using unbiased genome-scale screens consistently reveal new biology (Alvaro et al., 2007; Tkach et al., 2012). Indeed, our screens identified 105 genes with well-established roles in the DDR, and 240 genes with poorly understood or previously unappreciated phenotypes associated with DDR defects. There are two important features of our experimental pipeline that enabled discovery of this collection of potential new participants in the DDR. First, many genes were only linked to the DDR

by screening in chemically or genetically sensitized backgrounds, consistent with previous systematic exploration of genetic interactions causing growth defects (Costanzo et al., 2010), and highlighting the importance of the automated genetics component of our method. Second, automated image analysis facilitated accurate measurement of a detailed cell biological phenotype – in this case, the DNA damage focus – that provides a highly sensitive assay for defects in the DDR. Although defects in the DDR often translate into cell growth defects, we identified 164 mutants in our cell biological screens that were not identified using fitness based assays, either in standard growth conditions or in the presence of DNA damaging agents (Alvaro et al., 2007; Aouida et al., 2004; Begley et al., 2002; Bennett et al., 2001; Chang et al., 2002; Costanzo et al., 2010; Hartman and Tippery, 2004; Hillenmeyer et al., 2008; Parsons et al., 2004; Woolstencroft et al., 2006).

In an effort to understand DNA damage focus phenotypes that were only evident in a sensitized background, we focused on *VID22*, which was recently shown to be involved in the DDR, but whose relationship to *SGS1* was largely unexplored (Bonetti et al., 2012). Our experiments revealed shared roles for *VID22* and *SGS1* in minimally two facets of genome integrity maintenance. First, several of our phenotypic tests suggest a significant role for Sgs1 and Vid22 in rDNA integrity. Both unequal sister chromatid exchange and hyper-recombination of the rDNA repeats were elevated in *vid22*Δ *sgs1*Δ mutant populations, mirroring phenotypes see in Bloom syndrome, a disease caused by mutations in the mammalian homolog of *SGS1* (*BLM*; Grierson et al., 2013; Langlois et

al., 1989; Wang et al., 2003). Also, *vid22*Δ mutants were sensitive to DNA breaks that can only be repaired by NHEJ. These observations are consistent with previous work showing that Vid22 is required for recruitment of the DNA ligase Dnl4 to double-strand breaks, which is necessary for DNA repair by NHEJ (Bonetti et al., 2012; Grierson et al., 2013; Wilson et al., 1997). NHEJ is the preferred method of break repair at the rDNA locus (Torres-Rosell et al., 2007), and in the absence of *SGS1*, Dnl4 has a special role in double strand break repair as a result of collapsed replication forks at replication fork barriers in the rDNA cassette. Together, these observations implicate both Vid22 and Sgs1 in the repair of breaks via NHEJ, which is required for rDNA stability.

Our experiments also suggest a second role for Vid22 and Sgs1 in maintaining genome stability. We discovered an enrichment of Vid22 binding sites at predicted G4 DNA regions in the genome, implicating Vid22 in the processing, prevention or removal of RNA-DNA heteroduplex structures, which are associated with G4 DNA. Consistent with this possibility, for both *VID22* and genes encoding members of the RNase HII complex, which processes RNA-DNA hybrids, we observed strong negative genetic interactions with genes encoding components of the Smc5-6 complex, which is known to be required for the removal of X-shaped DNA structures that arise between sister chromatids during DNA repair (Bermudez-Lopez et al., 2010). These observations suggest that RNA-DNA hybrids may accumulate in the absence of either Vid22 or the RNase HII proteins, causing replication fork stalls and collapses that require the SMC5-6 complex to resolve. A role for Vid22 in dealing with RNA-DNA hybrids may also involve Sgs1, because this

helicase is important for removing the complicated DNA structures that result from collapsed replication forks, a phenotype that is exaggerated at the rDNA locus, which is especially prone to stalled and collapsed forks and elevated levels of RNA-DNA hybrids (Torres-Rosell et al., 2007). Interestingly, the mammalian homolog of *SGS1* (*BLM*) has been implicated in the unwinding of RNA-DNA hybrids (Grierson et al., 2013) and is a G4 helicase (Sun et al., 1999). Our automated imaging pipeline implemented in yeast cells may have identified a conserved pathway involving BED domain family proteins such as Vid22 and RecQ helicases (Sgs1-Blm1) in the maintenance of genome integrity through resolution of aberrant DNA structures linked to RNA-DNA hybrids.

So far, we have applied our integrated experimental and computational pipeline to the identification of defects in the DNA damage focus, by providing a cell biological readout of the DDR in response to thousands of genetic perturbations and an environmental insult. Our method is readily extensible to other fluorescent markers covering fundamental subcellular compartments or structures, as well as markers of important phenotypes, such as aging and cell death. Data from single mutant screens and other sources can then be used to select double mutant backgrounds and environmental perturbations for analysis with each marker. The functional redundancy built into eukaryotic cells means that genetic and environmental sensitization is required for a complete analysis. The resultant datasets will identify comprehensive sets of genes that influence the function of fundamental compartments and structures in a model eukaryotic cell. Our approach provides a proof-of-principle for both a technical and

conceptual platform that should be readily adapted to other cell types and more complex systems.

Experimental Procedures

High Throughput Preparation and Imaging of Yeast Cells

S. cerevisiae strains are listed in Table S1. Strains containing *RAD52-GFP, HTA2-mCherry and RPL39*pr-*tdTomato* (with *sgs1*Δ or *yku80*Δ in some experiments) were crossed to the deletion collection (Giaever et al. 2002) and to a collection of mutants carrying TS alleles of essential genes (Figure 1A,B; Li et al., 2011). Cells were prepared for imaging as previously described (Cox et al., 2015). All non-essential gene deletion mutants were grown at 30°C and all TS mutants in essential genes were grown at 22°C and then incubated for three hours at 37°C prior to imaging. Cells were imaged using an automated spinning disk confocal microscope (Evotec Opera™, PerkinElmer) as previously described (Chong et al., 2015).

Image Analysis

The segmentation pipeline used to identify fluorescently tagged cells is compatible with CellProfiler™ version 1.0.5811, and is available in full in the Supplemental Experimental Procedures. A support vector machine-based classifier was used to identify cells with Rad52-GFP foci and bootstrapping was used to determine the ideal cell count for reliable measurement of foci frequency (Supplemental Experimental Procedures).

Confirmations and Statistical Analysis

Genes identified in the imaging screen were retested using mutants from the Nat marked deletion collection (Costanzo et al., 2010) and/or by complementation with plasmids from the Molecular Barcoded Yeast (MoBY) ORF collection (Ho et al., 2009). Scoring enrichment and assessment of the relationship between foci score and various evolutionary / physiological features is described in the Supplemental Experimental Procedures.

Vid22 Experiments

Procedures exploring the biological role of Vid22 are described in detail in the Supplemental Experimental Procedures. Chromatin immunoprecipitation of Vid22 to double-strand breaks and ChIP-Seq assays were carried out as previously described (Ribeyre and Shore, 2012). Southern blot analysis was done as described (Medvedik and Sinclair, 2007). Calling card analysis was performed as described (Wang et al., 2011). Enrichment analysis of Vid22 binding sites identified in these assays is explained in detail in the Supplemental Experimental Procedures.

Instability at the rDNA locus was assessed by measuring rate of loss of an *ADE2* marker integrated into the rDNA (Kaeberlein et al., 1999) and sensitivity to galactose-inducible HO DSBs was assayed as described in Haber (2002). Genetic interactions and correlations with *VID22* were identified as described (Costanzo et al., 2010).

Author Contributions

B.J.A., C.B., B.L., M.M-F., C.L.M., R.D.M, D.S., G.W.B. and Z.Z. designed and supervised the project. E.B.S, K.J.F., T.L.S., D.A., C.R., V.R., D.M., D.N. and M.G carried out and analyzed experiments. E.B.S., K.J.F. and L.A.Z. performed large-scale analysis and interpretation. J.R., M.C., M.U., A.J.V. and E.N.K. performed additional data analysis. E.B.S., L.A.Z., T.L.S., D.A., E.N.K. and D.N. created figures. B.J.A., C.B., and E.B.S wrote the manuscript.

Acknowledgements

We thank Helena Friesen, Michael Cox, Yolanda Chong, Elena Kuzmin, Edith Cheng, Zhen-Yuan Lin and Anne-Claude Gingras for technical assistance and advice. We thank Harsha Garadi for comments on the manuscript. This work was supported by NSERC Discovery Grant RGPIN 326897-12 and Impact Grant 702310 from the Canadian Cancer Society Research Institute to G.W.B., and with funds from the Canadian Institute of Health Research to E.B.S. Infrastructure for high-content imaging was acquired with funds from the Canadian Foundation for Innovation and the Ministry of Research and Innovation (Ontario; grant 21745 to B.J.A. and C.B.). B.J.A. and C.B. are Senior Fellows in the Genetic Networks program of the Canadian Institute for Advanced Research.

References

Aguilera, A., and García-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. Mol. Cell *46*, 115–124.

Alvaro, D., Lisby, M., and Rothstein, R. (2007). Genome-wide analysis of Rad52 foci reveals diverse mechanisms impacting recombination. PLoS Genet. *3*, e228.

Aouida, M., Pagé, N., Leduc, A., Peter, M., and Ramotar, D. (2004). A genome-wide screen in Saccharomyces cerevisiae reveals altered transport as a mechanism of resistance to the anticancer drug bleomycin. Cancer Res. *64*, 1102–1109.

Aravind, L. (2000). The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. Trends Biochem. Sci. *25*, 421–423.

Askree, S.H., Yehuda, T., Smolikov, S., Gurevich, R., Hawk, J., Coker, C., Krauskopf, A., Kupiec, M., and McEachern, M.J. (2004). A genome-wide screen for Saccharomyces cerevisiae deletion mutants that affect telomere length. Proc. Natl. Acad. Sci. USA *101*, 8658–8663.

Badis, G., Chan, E.T., van Bakel, H., Peña-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L., et al. (2008). A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters. Mol. Cell *32*, 878–887.

Baryshnikova, A., Costanzo, M., Kim, Y., Ding, H., Koh, J., Toufighi, K., Youn, J.-Y., Ou, J., San Luis, B.-J., Bandyopadhyay, S., et al. (2010). Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. Nat. Methods *7*, 1017–1024.

Begley, T.J., Rosenbach, A.S., Ideker, T., and Samson, L.D. (2002). Damage recovery pathways in Saccharomyces cerevisiae revealed by genomic phenotyping and interactome mapping. Mol. Cancer Res. *1*, 103–112.

Bennett, C.B., Lewis, L.K., Karthikeyan, G., Lobachev, K.S., Jin, Y.H., Sterling, J.F., Snipe, J.R., and Resnick, M.A. (2001). Genes required for ionizing radiation resistance in yeast. Nat. Genet. *29*, 426–434.

Bermudez-Lopez, M., Ceschia, A., de Piccoli, G., Colomina, N., Pasero, P., Aragon, L., and Torres-Rosell, J. (2010). The Smc5/6 complex is required for dissolution of DNA-mediated sister chromatid linkages. Nucleic Acids Res. *38*, 6502–6512.

Bonetti, D., Anbalagan, S., Lucchini, G., Clerici, M., and Longhese, M.P. (2012). Tbf1 and Vid22 promote resection and non-homologous end joining of DNA double-strand break ends. EMBO J 1–15.

Breker, M., Gymrek, M., and Schuldiner, M. (2013). A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. J. Cell. Biol. *200*, 839–850.

Brown, C.R., Cui, D.Y., Hung, G.G., and Chiang, H.L. (2001). Cyclophilin A mediates Vid22p function in the import of fructose-1,6-bisphosphatase into Vid vesicles. J. Biol. Chem. *276*, 48017–48026.

Brown, C.R., McCann, J.A., Hung, G.G.-C., Elco, C.P., and Chiang, H.-L. (2002). Vid22p, a novel

plasma membrane protein, is required for the fructose-1,6-bisphosphatase degradation pathway. J. Cell. Sci. *115*, 655–666.

Capra, J.A., Paeschke, K., Singh, M., and Zakian, V.A. (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae. PLoS Comput. Biol. *6*, e1000861.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. *7*, R100.

Chan, Y.A., Aristizabal, M.J., Lu, P.Y.T., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip. PLoS Genet. *10*, e1004288.

Chang, M., Bellaoui, M., Boone, C., and Brown, G. (2002). A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. Proc. Natl. Acad. Sci. USA *99*, 16934–16939.

Chong, Y.T., Koh, J.L.Y., Friesen, H., Duffy, K., Cox, M.J., Moses, A., Moffat, J., Boone, C., and Andrews, B.J. (2015). Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. Cell *161*, 1413–1424.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., 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 *446*, 806–810.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S., et al. (2010). The Genetic Landscape of a Cell. Science *327*, 425–431.

Cox, M.J., Chong, Y.T., Boone, C., and Andrews, B. (2015). Chapter 15; Liquid Growth of Arrayed Fluorescently Tagged *Saccharomyces cerevisiae* Strains for Live-Cell High-Throughput Microscopy Screens. In Budding Yeast: a Laboratory Manual, B. Andrews, C. Boone, T.N. Davis, and S. Fields, eds. (Cold Spring Harbor Laboratory Press).

Elston, R.C. (1991). On Fisher's Method of Combining p-Values. Biomedical Journal *33*, 339–345.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Grierson, P.M., Acharya, S., and Groden, J. (2013). Collaborating functions of BLM and DNA topoisomerase I in regulating human rDNA transcription. Mutat. Res. *743-744*, 89–96.

Haoyi Wang, David Mayhew, Chen, X., Mark Johnston, and Robi David Mitra (2011). Calling Cards enable multiplexed identification of the genomic targets of DNA-binding proteins. Genome Res. *21*, 748–755.

Hartman, J.L., and Tippery, N.P. (2004). Systematic quantification of gene interactions by phenotypic array analysis. Genome Biol. *5*, R49.

Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St Onge, R.P., Tyers, M., Koller, D., et al. (2008). The Chemical Genomic Portrait of Yeast:

Uncovering a Phenotype for All Genes. Science *320*, 362–365.

Ho, C.H., Magtanong, L., Barker, S.L., Gresham, D., Nishimura, S., Natarajan, P., Koh, J.L.Y., Porter, J., Gray, C.A., Andersen, R.J., et al. (2009). A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. Nat. Biotechnol. *27*, 369–377.

Huh, W.-K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686–691.

Kaeberlein, M., McVey, M., and Guarente, L. (1999). The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms. Genes Dev. *13*, 2570–2580.

Kofoed, M., Milbury, K.L., Chiang, J.H., Sinha, S., Ben-Aroya, S., Giaever, G., Nislow, C., Hieter, P., and Stirling, P.C. (2015). An Updated Collection of Sequence Barcoded Temperature-Sensitive Alleles of Yeast Essential Genes. G3 *5*:1879-87.

Langlois, R.G., Bigbee, W.L., Jensen, R.H., and German, J. (1989). Evidence for increased in vivo mutation and somatic recombination in Bloom's syndrome. Proc. Natl. Acad. Sci. USA *86*, 670–674.

Levy, S.F., and Siegal, M.L. (2008). Network hubs buffer environmental variation in Saccharomyces cerevisiae. PLoS Biol. *6*, e264.

Li, Z., Vizeacoumar, F.J., Bahr, S., Li, J., Warringer, J., S Vizeacoumar, F., Min, R., VanderSluis, B., Bellay, J., Devit, M., et al. (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. Nat. Biotechnol. *29*, 361–367.

Lisby, M., Rothstein, R., and Mortensen, U.H. (2001). Rad52 forms DNA repair and recombination centers during S phase. Proc. Natl. Acad. Sci. USA *98*, 8276–8282.

Lisby, M., Barlow, J.H., Burgess, R.C., and Rothstein, R. (2004). Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. Cell *118*, 699–713.

Lisby, M., Mortensen, U.H., and Rothstein, R. (2003). Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat. Cell. Biol. *5*, 572–577.

Lukas, C., Savic, V., Bekker-Jensen, S., Doil, C., Neumann, B., Pedersen, R.S., Grøfte, M., Chan, K.L., Hickson, I.D., Bartek, J., et al. (2011). 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat. Cell. Biol. *13*, 243–253.

Maizels, N., and Gray, L.T. (2013). The G4 Genome. PLoS Genet. *9*:e1003468.

Malo, N., Hanley, J., Cerquozzi, S., Pelletier, J., and Nadon, R. (2006). Statistical practice in high-throughput screening data analysis. Nat. Biotechnol. *24*, 167–175.

Medvedik, O., and Sinclair, D.A. (2007). Caloric restriction and life span determination of yeast cells. Methods Mol. Biol. *371*, 97–109.

Mnaimneh, S., Davierwala, A.P., Haynes, J., Moffat, J., Peng, W.-T., Zhang, W., Yang, X., Pootoolal, J., Chua, G., Lopez, A., et al. (2004). Exploration of Essential Gene Functions via Titratable Promoter Alleles. Cell *118*, 31–44.

Ohya, Y., Sese, J., Yukawa, M., Sano, F., Nakatani, Y., Saito, T.L., Saka, A., Fukuda, T., Ishihara, S., Oka, S., et al. (2005). High-dimensional and large-scale phenotyping of yeast mutants. Proc. Natl. Acad. Sci. USA *102*, 19015–19020.

Parsons, A., Brost, R., Ding, H., Li, Z., and Zhang, C. (2004). Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. Nat. Biotechnol. *22*, 62-9.

Preti, M., Ribeyre, C., Pascali, C., Bosio, M.C., Cortelazzi, B., Rougemont, J., Guarnera, E., Naef, F., Shore, D., and Dieci, G. (2010). The telomere-binding protein Tbf1 demarcates snoRNA gene promoters in Saccharomyces cerevisiae. Mol. Cell *38*, 614–620.

Ribaud, V., Ribeyre, C., Damay, P., and Shore, D. (2012). DNA-end capping by the budding yeast transcription factor and subtelomeric binding protein Tbf1. EMBO J *31*, 138–149.

Ribeyre, C., and Shore, D. (2012). Anticheckpoint pathways at telomeres in yeast. Nat. Struct. Mol. Biol. *19*, 307–313

Skellam, J. (1948). A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. Journal of the Royal Statistical Society, Series B (Methodological) *10*, 257–261.

Sun, H., Bennett, R.J., and Maizels, N. (1999). The Saccharomyces cerevisiae Sgs1 helicase efficiently unwinds G-G paired DNAs. Nucleic Acids Res. *27*, 1978–1984.

Tkach, J.M., Yimit, A., Lee, A.Y., Riffle, M., Costanzo, M., Jaschob, D., Hendry, J.A., Ou, J., Moffat, J., Boone, C., et al. (2012). Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. Nat. Cell Biol. *14*, 966–976.

Tong, A.H.Y., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W.V., Bussey, H., et al. (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364–2368.

Torres-Rosell, J., Machin, F., Farmer, S., Jarmuz, A., Eydmann, T., Dalgaard, J.Z., and Aragon, L. (2005). SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. Nat. Cell Biol. *7*, 412–419.

Torres-Rosell, J., Sunjevaric, I., De Piccoli, G., Sacher, M., Eckert-Boulet, N., Reid, R., Jentsch, S., Rothstein, R., Aragon, L., and Lisby, M. (2007). The Smc5-Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. Nat. Cell Biol. *9*, 923–931.

Vizeacoumar, F.J., van Dyk, N., S Vizeacoumar, F., Cheung, V., Li, J., Sydorskyy, Y., Case, N., Li, Z., Datti, A., Nislow, C., et al. (2010). Integrating high-throughput genetic interaction mapping and high-content screening to explore yeast spindle morphogenesis. J. Cell Biol. *188*, 69–81.

Wang, H., Mayhew, D., Chen, X., Johnston, M., and Mitra, R.D. (2011). Calling Cards enable multiplexed identification of the genomic targets of DNA-binding proteins. Genome Res. *21*, 748–755.

Wang, W., Seki, M., Narita, Y., Nakagawa, T., Yoshimura, A., Otsuki, M., Kawabe, Y.I., Tada, S., Yagi, H., Ishii, Y., et al. (2003). Functional Relation among RecQ Family Helicases RecQL1, RecQL5, and BLM in Cell Growth and Sister Chromatid Exchange Formation. Mol.Cell. Bio. *23*, 3527–3535.

Wilson, T.E., Grawunder, U., and Lieber, M.R. (1997). Yeast DNA ligase IV mediates non-homologous DNA end joining. Nature *388*, 495–498.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901–906.

Woolstencroft, R.N., Beilharz, T.H., Cook, M.A., Preiss, T., Durocher, D., and Tyers, M. (2006). Ccr4 contributes to tolerance of replication stress through control of CRT1 mRNA poly(A) tail length. J. Cell Sci. *119*, 5178–5192.

Figure Legends

Figure 1. Synthetic Genetic Array-High Content Screening (SGA-HCS) Strategy for Identifying Cell Populations with Elevated Levels of DNA Damage Foci. See also Figure S1.

A. Diagram illustrating array construction strategy for automated image analysis of fluorescent proteins marking specific compartments within the cell. A *RAD52-GFP* fusion gene product marks DNA damage foci (green dot) while nuclear (*HTA2-mCherry*; dark red) and cytoplasmic (*RPL39*pr-*tdTomato*; light red) signals provide spatial and cell cycle context. A sensitizing gene deletion can be introduced into the query strain at this stage. The Synthetic Genetic Array method is used to introduce reporters and mutations of interest into the essential TS mutant and non-essential gene deletion collections via automated replica-pinning.
B. High-throughput (HTP) preparation of cells for automated imaging. Cells are transferred to liquid medium, or liquid medium containing drug to provide a chemically sensitized background. Objects in micrographs are segmented in CellProfiler™, and an SVM-based classification is used to separate cells that contain a DNA damage-induced focus from those that do not.
C. Illustration of strategy for identifying hits in SGA-HCS screens of DNA damage foci. A distribution is displayed in which the average frequency of foci in all single gene deletion and TS mutant populations across 11 biological replicates (n = $4.8x10^4$) is scored, and the wild-type average distribution is highlighted in gray. Five positive controls are indicated with tick marks in the outlier set.
D. Bootstrapping approach to select an optimal minimum cell count for analysis. The black dashed line indicates the standard deviation in foci levels at the selected sample size minimum (1000 cells/mutant indicated with dashed green line, standard deviation = 0.82%).
E. Graph illustrating fraction of mutant strains for which at least 1000 cells were imaged. SM = Single mutant non-essential deletion mutants; SM-TS = single mutant TS alleles of essential genes; SM+Phleo = non-essential deletion mutants with phleomycin; SM+Phleo-TS = TS allele array plus phleomycin; *yku80*Δ = non-essential deletion mutants lacking *YKU80*; *yku80*Δ-TS = TS allele array lacking *YKU80*; *sgs1*Δ = non-essential deletion mutants lacking *SGS1*; *sgs1*Δ-TS = TS allele array lacking *SGS1*.
F. Graph showing precision of five scoring methods. Single mutants were scored for frequency of DNA damage foci using several methods. Precision was scored on ranked mutants using as a standard all genes annotated to the DNA replication / repair / cohesion functional category in Costanzo et al., 2010. BD = Binomial distribution, F = Fisher's Score.

Figure 2. Mutants with Elevated Levels of Rad52-GFP Foci. See also Figures S2 and S3 and Tables S2 and S3.

Summary network of mutants with elevated levels of Rad52-GFP foci. The network diagram summarizes the results of all screens performed. Hub nodes indicate the screening condition (Single Mutants = BY4879; Phleomycin-treated = BY4879; *sgs1*Δ = BY4880; *yku80*Δ = BY4881), and edges connect these conditions to the hit genes whose deletion or conditional mutation is implicated in an elevated DNA damage foci phenotype. All hit nodes are color-coded according to functional category (legend below network), and those that confirmed in a secondary assay are outlined in black. Total hits = 345.

Figure 3. Functional Enrichments in Screens of Non-essential Gene Deletion Mutants for Elevated Levels of DNA damage Foci. See also Figure S4 and Table S4.

A. Functional enrichment of hits from SGA-HCS screens of single deletion mutant array. Right panel; functional categories are derived from (Costanzo et al., 2010), and are listed to the left of the heat map. Yellow indicates a positive log odds ratio (LOD), or enrichment, and blue indicates a negative LOD, or an underrepresentation (scale bar between panels a and b). KT = kinetochore, MT = microtubules, HR = homologous recombination. Left panel; Enrichment of hits in 16 genome-wide datasets, each assessing an aspect of the DNA damage response pathway. Yellow indicates a positive LOD ratio, or enrichment, and blue indicates a negative LOD, or an underrepresentation. MMS = methyl methanesulfonate; IR = ionizing radiation; HU = hydroxyurea; MDR = multidrug resistance genes, in both homozygous and heterozygous deletion sets.
B. Overlay of mutants exhibiting elevated levels of foci onto the yeast genetic interaction correlation network (Costanzo et al., 2010). The genetic interaction network described in Costanzo et al., 2010 is shown with the locations of 18 prominent bioprocess annotations outlined (solid lines). Non-essential genes identified as hits in our screens are overlaid on this network (green nodes), and the five bioprocesses in which these hits were most highly enriched in (Costanzo et al., 2010) are annotated in bold, with LOD and P-values indicated in italics (black outlines).

Figure 4: Localization of Vid22 to an Induced DSB and DNA Damage Focus Kinetics in the Absence of *VID22* and *SGS1*. See also Figures S5 and S6.

A. Schematic of a strain designed to query Vid22-Myc recruitment to an induced *HO* break. An *HO* cut site is integrated to the left of the centromere on Chromosome VII, which is acted upon by the galactose-inducible *HO* endonuclease. Two probes (Amp7 = blue; Amp14 = yellow) adjacent to the DSB are used to assess Vid22-Myc binding.

B. Southern blot analysis indicating HO endonuclease efficiently cleaves an integrated HO cut site. DNA from strains carrying a unique cut site for the HO endonuclease with (wild-type, BY5495; or *sgs1*Δ, BY5496) or without (no HO, BY5508) an integrated *GAL-HO* gene was digested with *Eco*RV. The blot was probed with both a ^{32}P-radiolabeled *ADE2* DNA fragment and a ^{32}P-radiolabeled *NMD5* fragment, and the uncut DNA, cut DNA and an internal control (*SNR52*) are indicated.
C. Chromatin immunoprecipitation (ChIP) of Vid22 to an HO-induced DNA double-stranded break. Vid22 recruitment was assessed using probes to two sites (Amp7 = blue; Amp14 = yellow), by ChIP of Vid22-Myc before (0 on X axis) and after (2 on X axis) induction of HO endonuclease. A Vid22-Myc strain in the presence (wild-type; BY5495) of absence of *SGS1* (*sgs1*Δ; BY5496) was cross-linked in formaldehyde, cells were lysed and Vid22-Myc was immunoprecipitated using a Myc-specific antibody. A strain with no HO site was used as a control (top panel; BY5508). Error bars represent the standard deviation between three replicate qPCR reactions.
D. Kinetic analysis of Rad52-GFP focus formation. Wild-type (BY4879), *vid22*Δ (BY5418), *sgs1*Δ (BY4880), and *vid22*Δ *sgs1*Δ (BY5433) cells expressing Rad52-GFP (green) and Hta2-RFP (red) in logarithmic growth phase were imaged every 30 min for 8 hrs. Merged projections of the DIC, green, and red channels are shown for the 0, 4 and 8 hour time-points, and representative cells containing DNA damage foci are highlighted in the right panel with white arrowheads.
E. Quantification of the number of Rad52-GFP foci per cell. Spontaneous Rad52 foci were counted in 100 wild-type, *vid22*Δ, *sgs1*Δ, and *vid22*Δ *sgs1*Δ cells as indicated (1 focus per cell = light green, 2 foci per cell = darker green, >2 foci per cell = darkest green).
F. Quantification of the duration of Rad52-GFP foci. Spontaneous Rad52-GFP foci were assessed every 30 minutes for more than 3 hours as indicated (X axis). Foci were followed in 100 wild-type (blue), *vid22*Δ (green), *sgs1*Δ (red), and *vid22*Δ *sgs1*Δ (yellow) cells. The percentage of cells with a persistent focus at each time point is shown.

Figure 5. Localization of Vid22 to Gene Promoters. See also Table S5.

A. Summary network of all loci identified as Vid22 binding sites by ChIP-Seq (BY5493) and calling card (BY5487) analyses. Green nodes indicate a ChIP or Calling card site that overlaps a region of G4 DNA (Capra et al., 2010), and black nodes represent those that do not overlap G4 regions.
B. Overlap of Vid22 binding sites identified by calling card analysis with Vid22, Env11 and Tbf1 ChIP-Seq analysis (Vid22 ChIP-Seq = blue, BY5493; Env11 ChIP-Seq = yellow, BY5494; Tbf1 ChIP-Seq = green, BY5507; Preti et al., 2010).
C. Effect of *SGS1* deletion on Vid22 recruitment to promoter regions. Association of Vid22-Myc with promoters of known target genes (*PAF1*, *MDM31*, *TBF1*) and a negative control gene (*NME1*) was assessed using ChIP as described in the

legend of Figure 4 (Wild-type = black; *sgs1*Δ = green). Error bars represent the standard deviation between three replicate qPCR reactions.

D. Enrichment of Vid22 binding sites at regions that overlap G4 DNA structures. Fold enrichment over background of Vid22 binding at predicted G4 DNA regions in Vid22 calling card (green) and ChIP-Seq data (black) is shown (* = *P-value* < 0.03; ** = *P-value* < $1.7x10^{-13}$). Less significant enrichments at regions that are predisposed to elevated levels of RNA-DNA heteroduplex formation in wild-type (RNA-DNA hybrids wild-type) and an RNase HI and HII mutant strain (rnh1Δ rnh201Δ; Chan et al., 2014). Vid22 ChIP-Seq = black; Vid22 calling card = green.

Figure 6: Two-dimensional Hierarchical Clustering of Synthetic Genetic Interactions Associated with *VID22.*

A. Component of a large cluster-gram of genetic interactions involving deletion mutants of non-essential genes and TS allele mutants of essential genes (unpublished data available at http://andrewslab.ccbr.utoronto.ca/supplement/styles2015/; Username = styles2015, Password = microscopy; Costanzo et al., 2010). Array genes (X-axis) and query genes (Y-axis) are hierarchically clustered based on genetic interaction score (yellow = positive GI, blue = negative GI, black = no GI; Baryshnikova et al., 2010). Upper case gene names indicate non-essential genes screened as deletion mutants (*VID22* in bold), and lower case gene names are associated with TS alleles of essential genes (different alleles are indicated by a unique allele number or designation).
B. Association of Rnh202-Myc (top; BY5501), Rnh201-Myc (middle; BY5498) and Rnh203-Myc (bottom; BY5504) with a DNA double strand break site. ChIP was performed after zero and two-hour induction of *HO* endonuclease as described in the legend of Figure 4. RNH recruitment was assessed using probes to two sites (Amp7 = blue; Amp14 = yellow). Error bars represent the standard deviation between three replicate qPCR reactions.

Figure 1

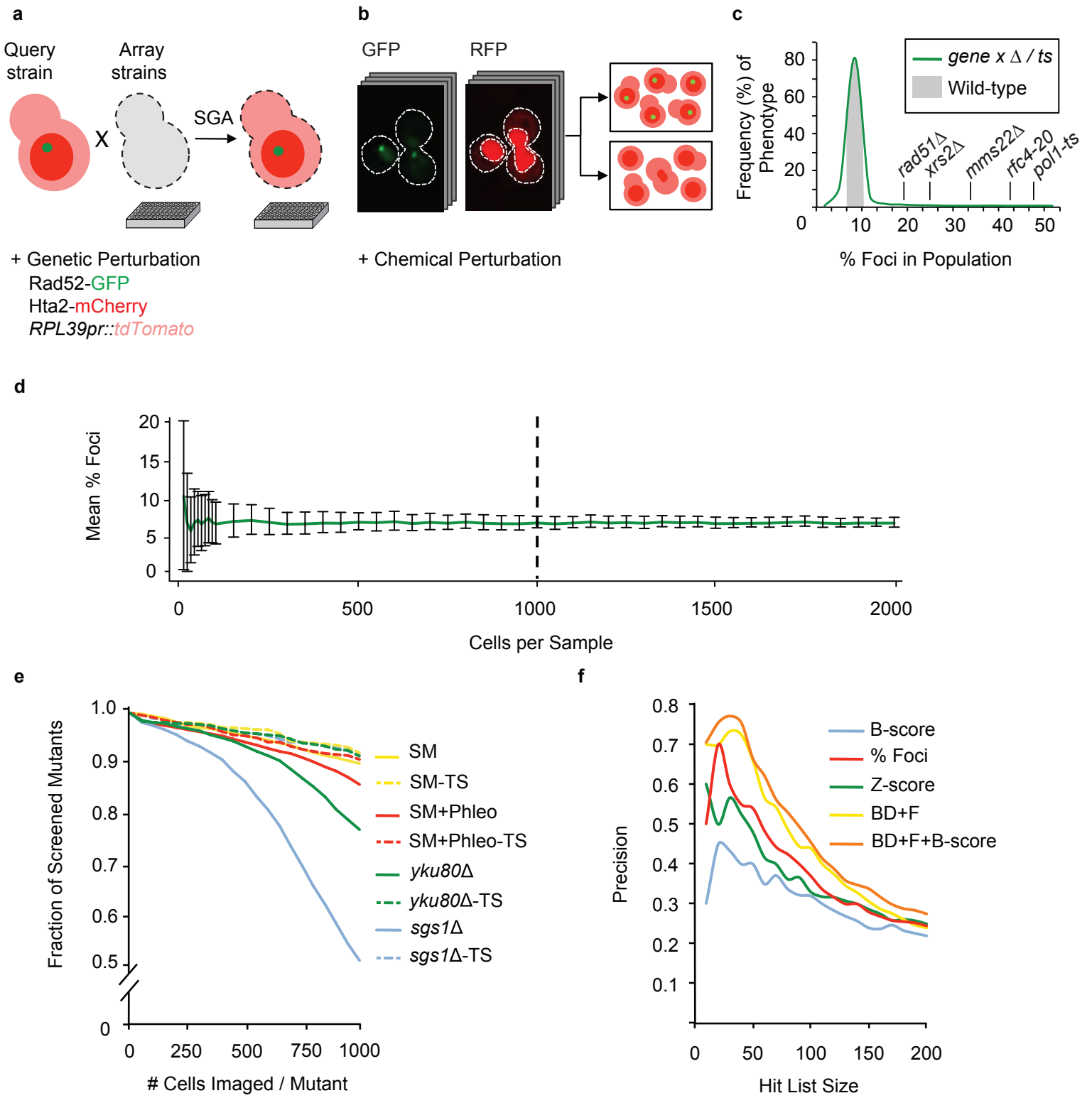


Figure 1: Synthetic Genetic Array-High Content Screening (SGA-HCS) strategy for identifying cell populations with elevated levels of DNA damage foci.

Figure 2

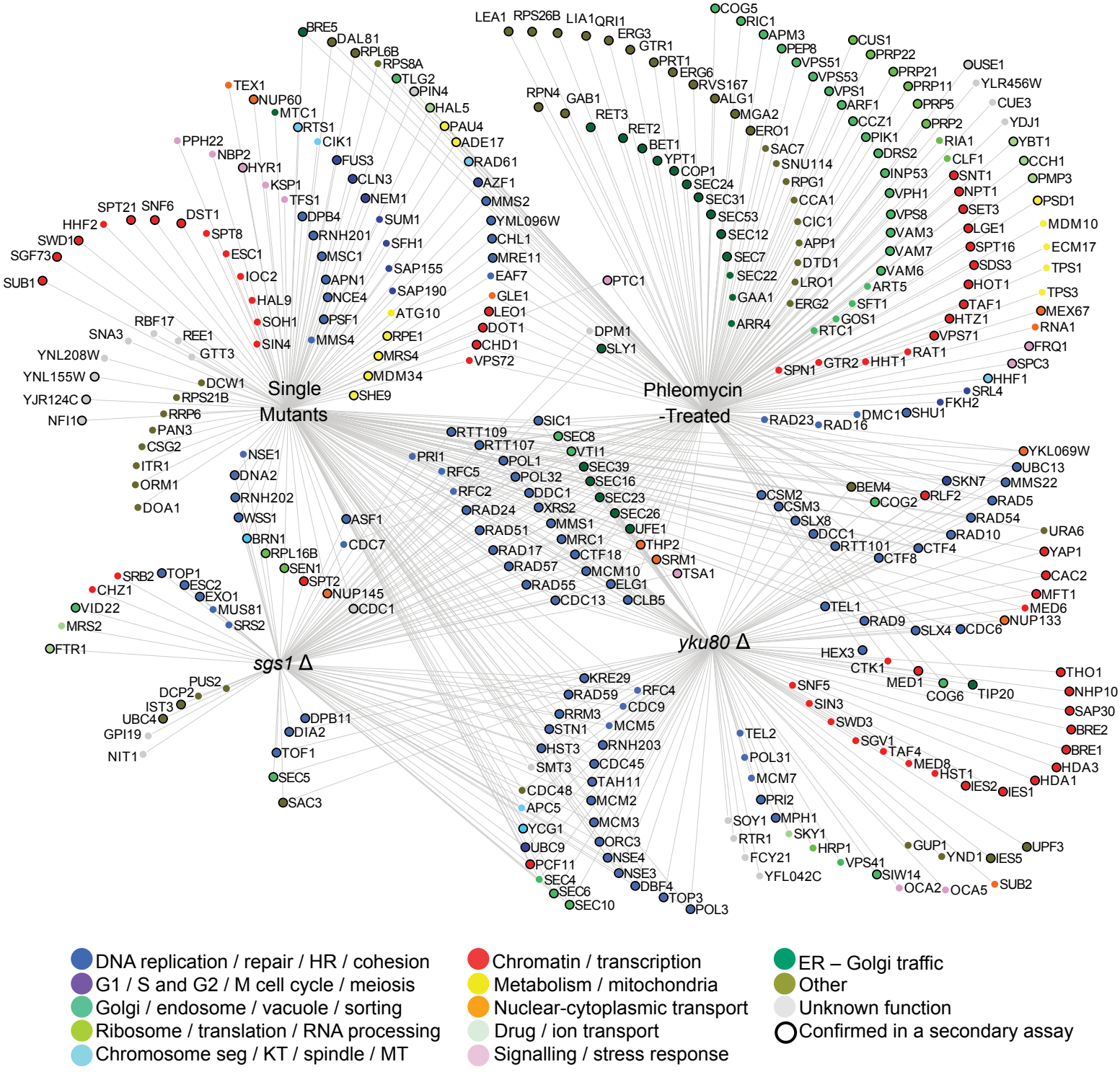
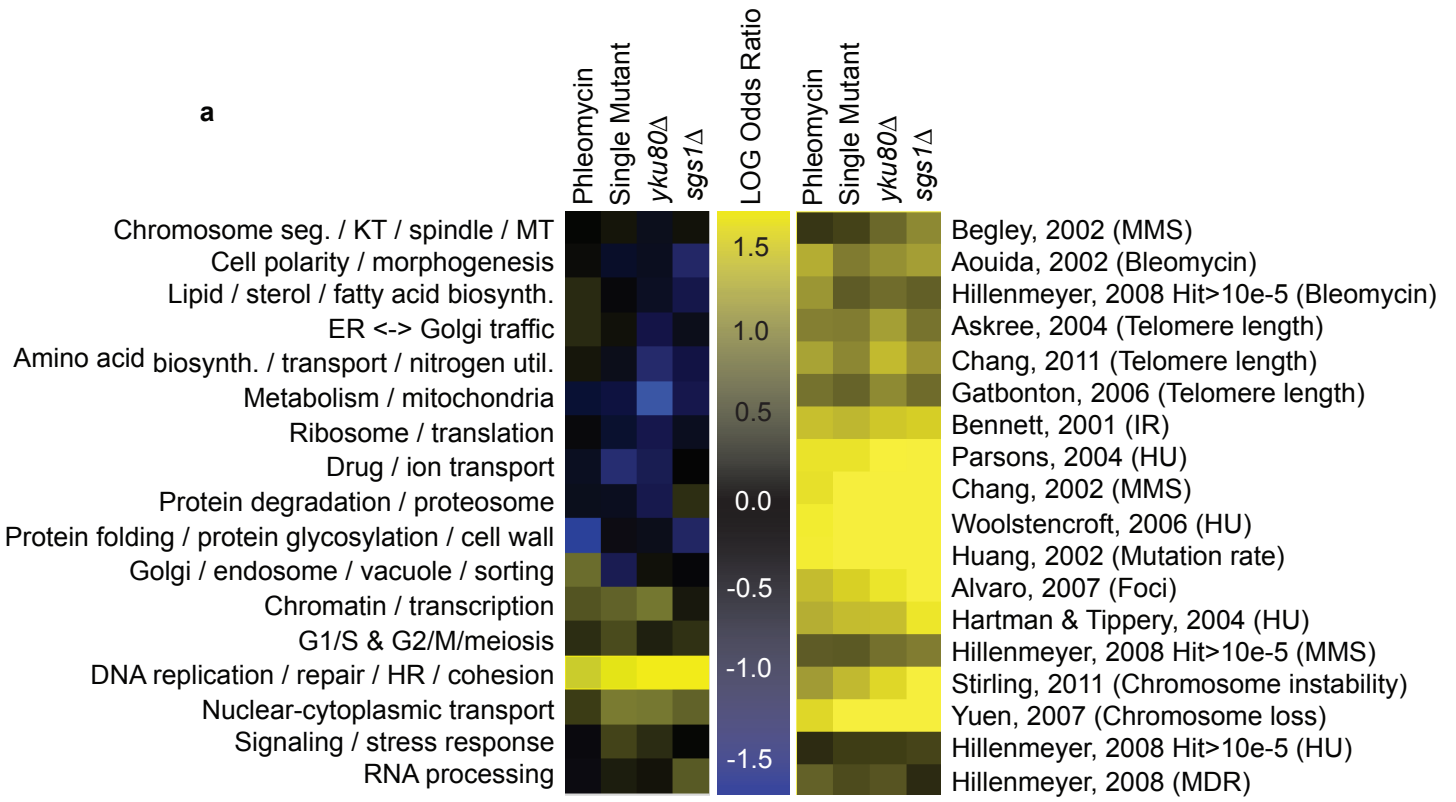


Figure 2: Mutants with elevated levels of Rad52-GFP foci.

Figure 3



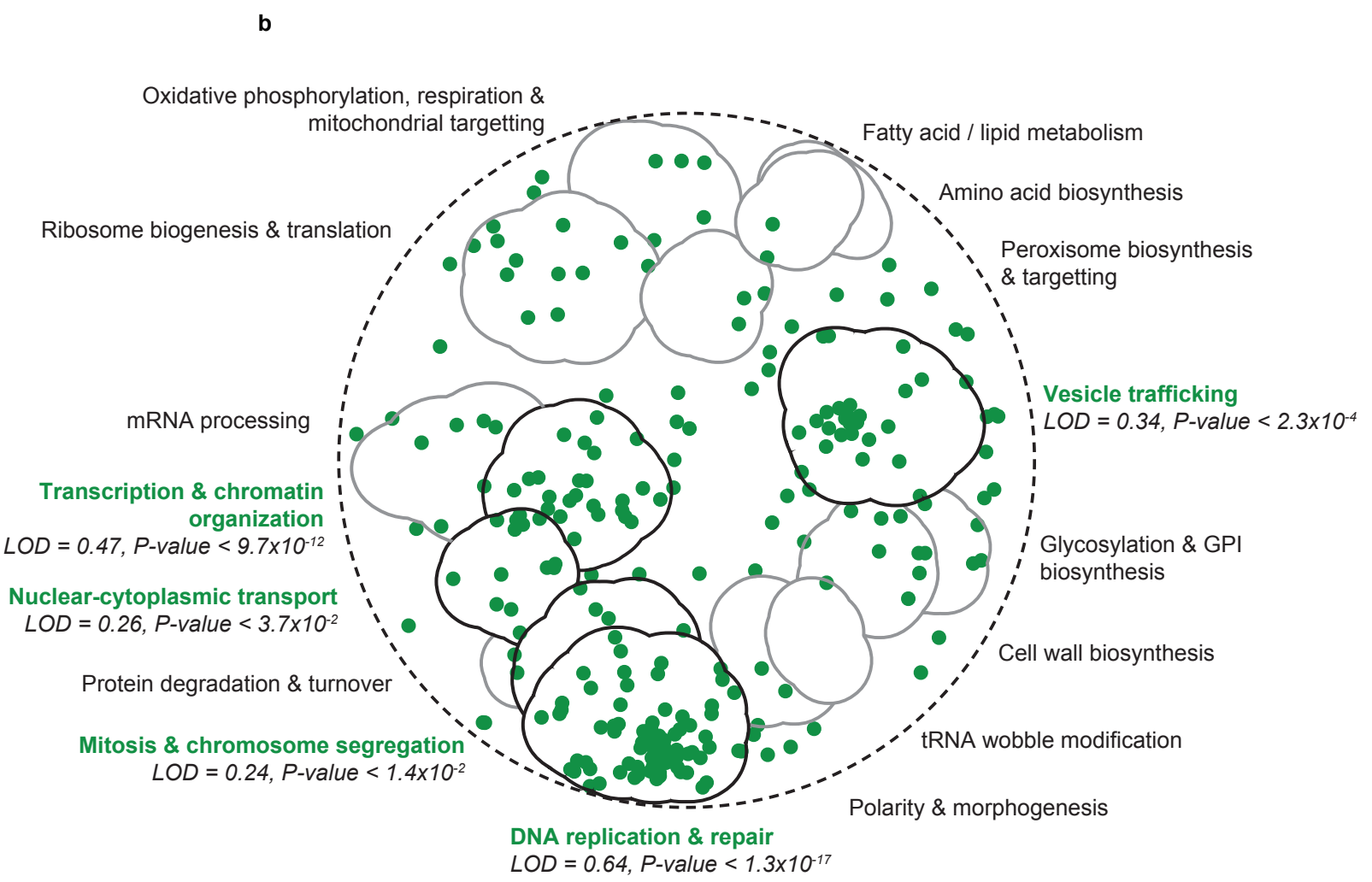


Figure 3: Functional enrichments in screens of non-essential gene deletion mutants for elevated levels of DNA damage foci.

Figure 4

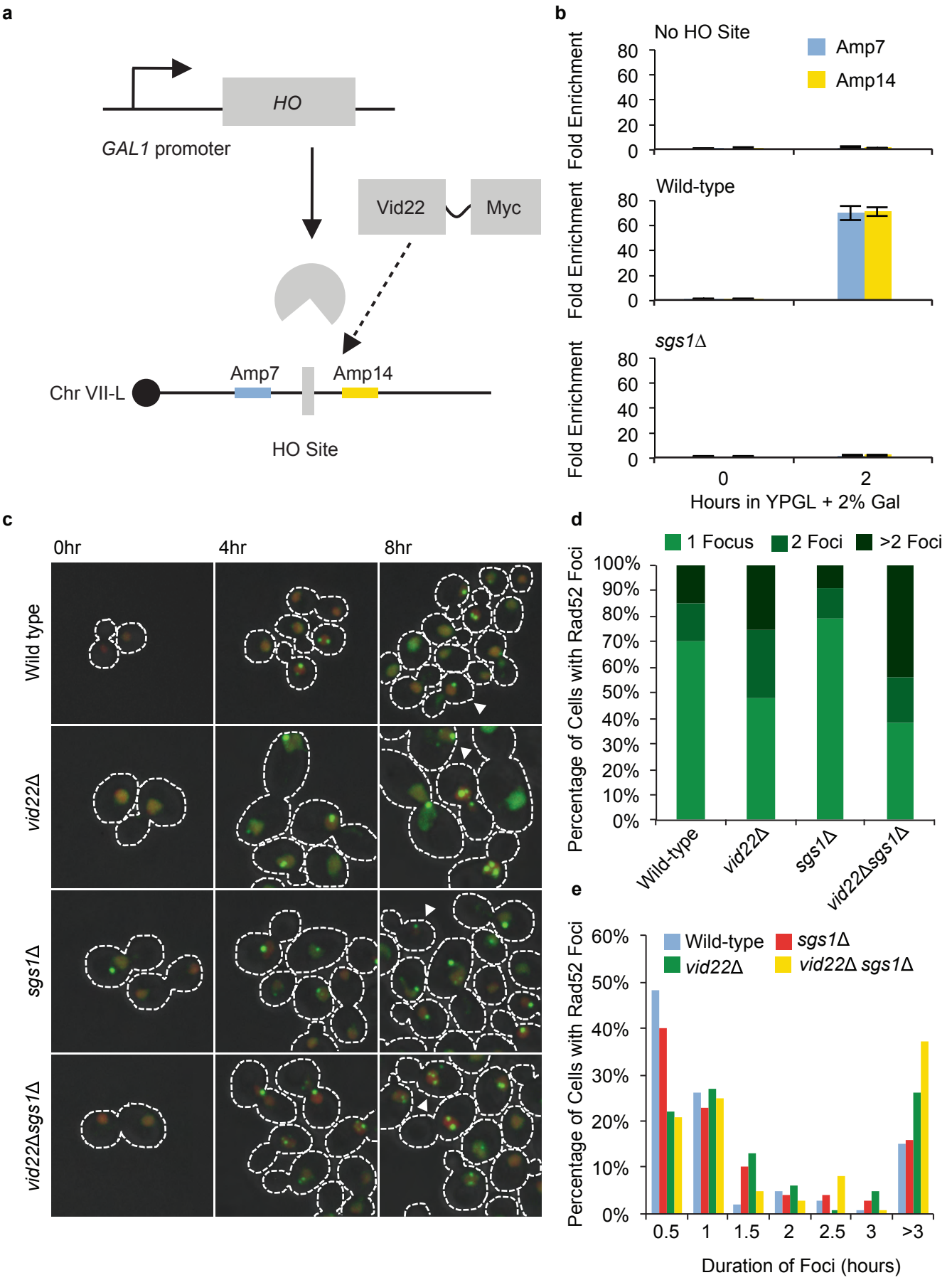


Figure 4: Localization of Vid22 to an induced DSB and DNA damage focus kinetics in the absence of *VID22* and *SGS1*.

Figure 5

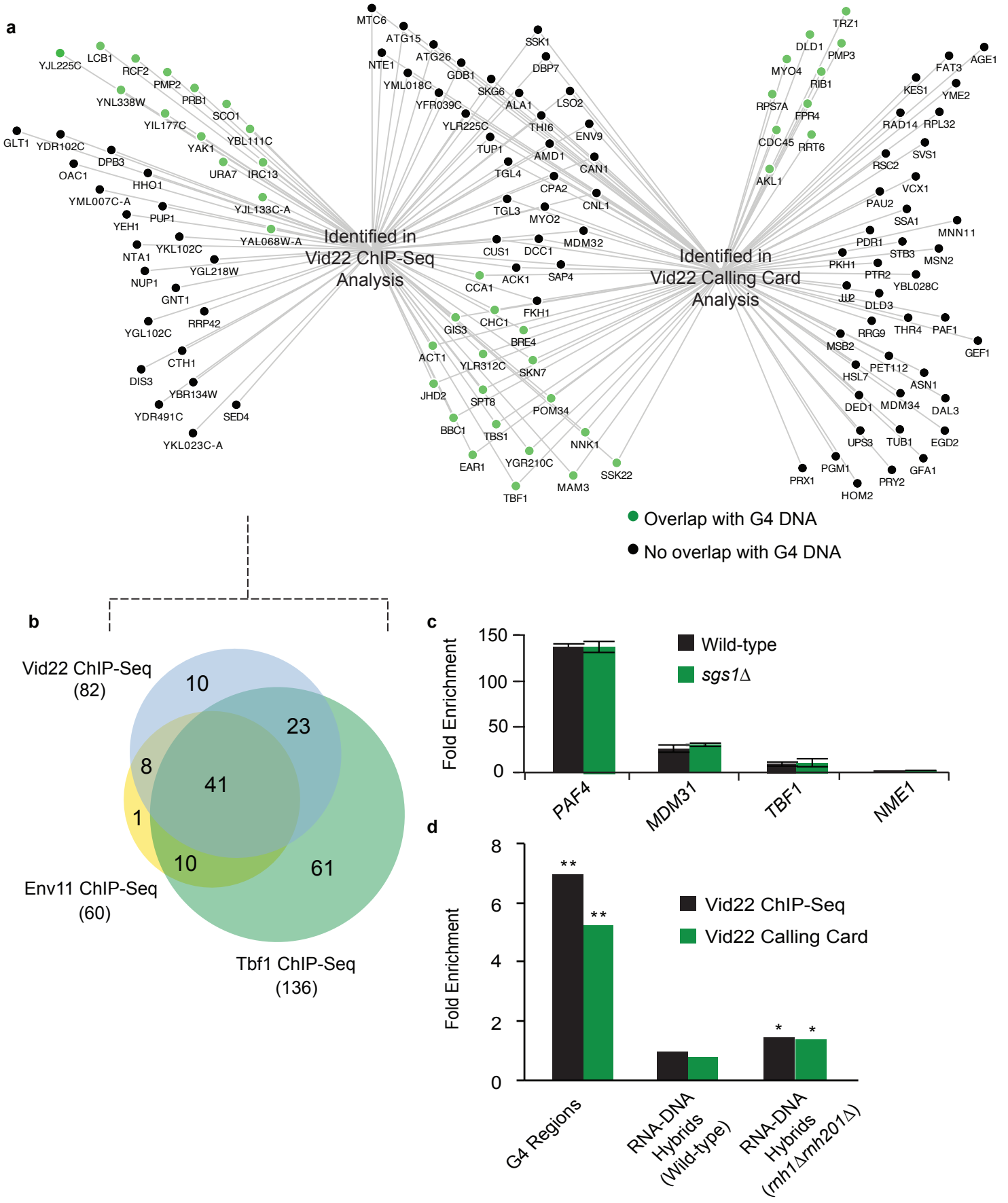


Figure 5: Localization of Vid22 to gene promoters.

Figure 6

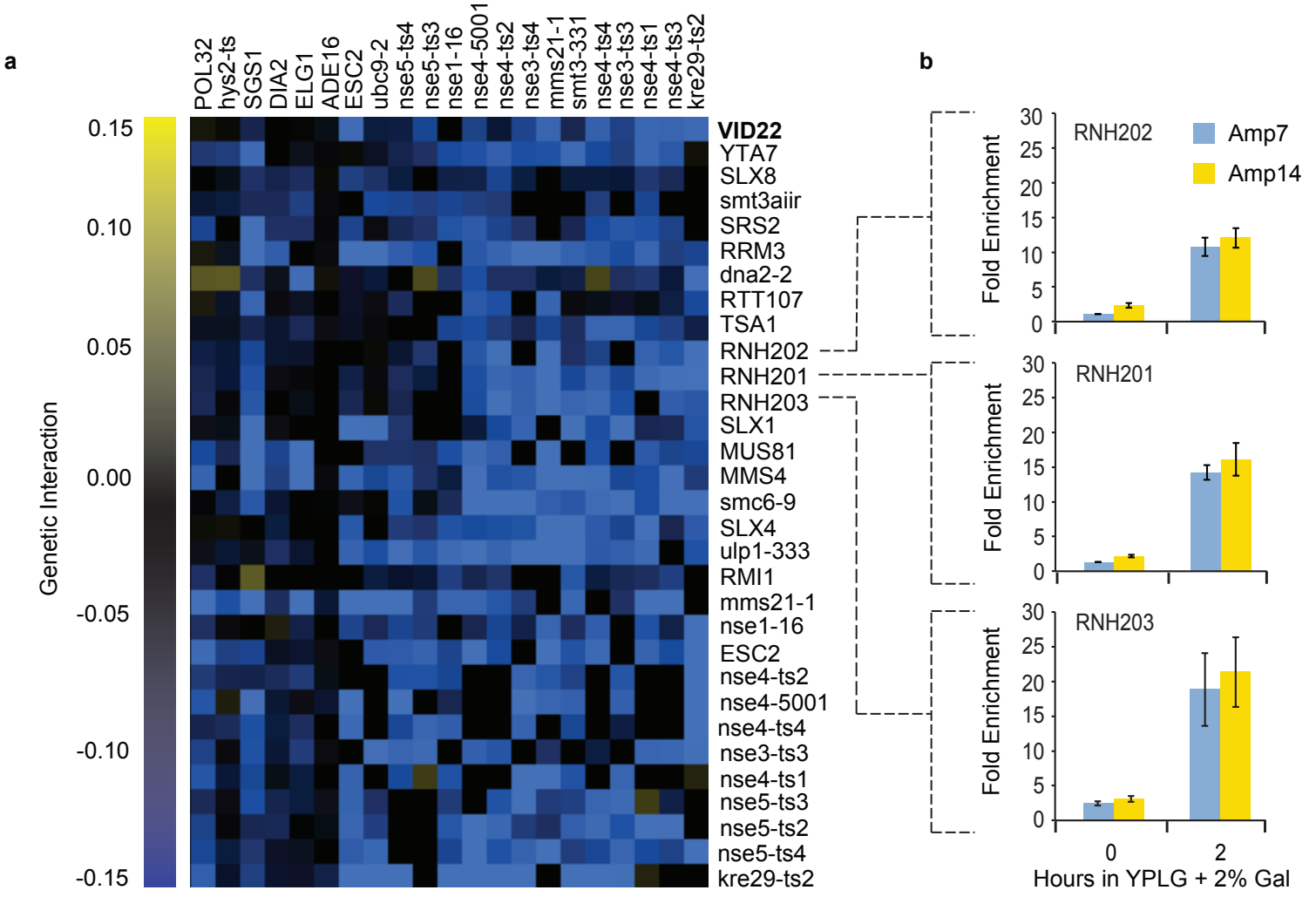


Figure 6: Two-dimensional hierarchical clustering of synthetic genetic interactions associated with *VID22.*

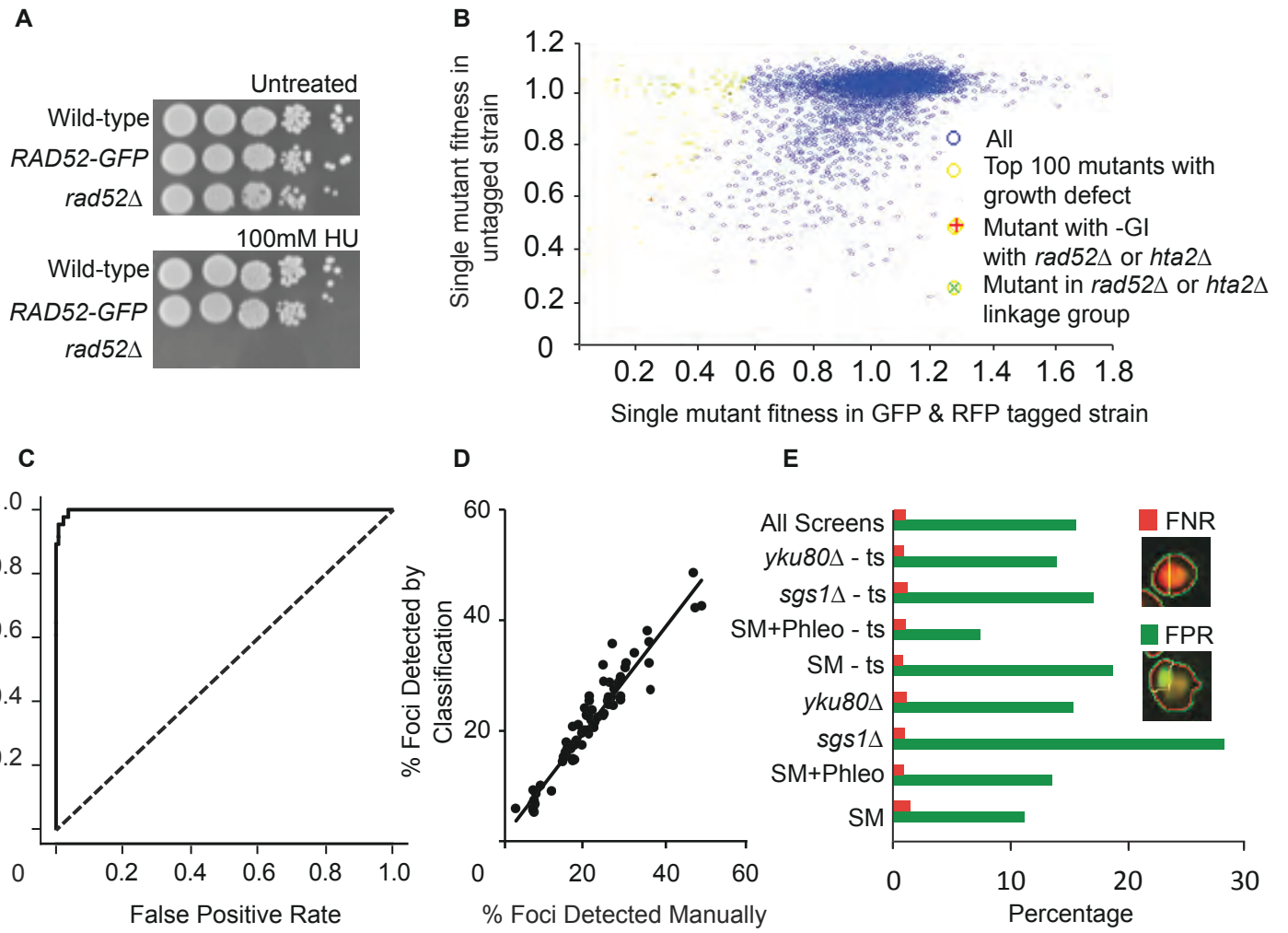
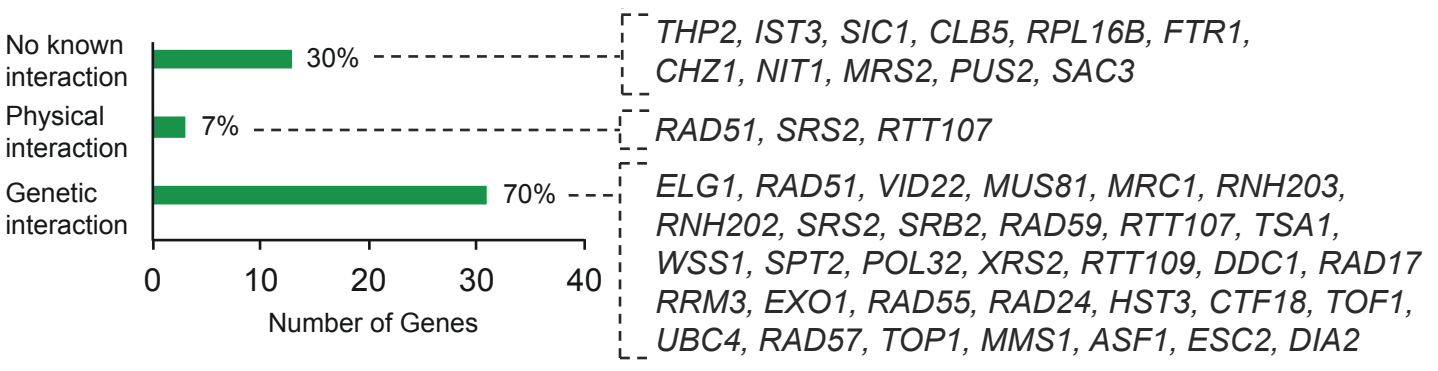

A
Untreated
Wild-type
RAD52-GFP
rad52Δ
100mM HU
Wild-type
RAD52-GFP
rad52Δ
B
Single mutant fitness in untagged strain
Single mutant fitness in GFP & RFP tagged strain
All
Top 100 mutants with growth defect
Mutant with -GI with rad52Δ or hta2Δ
Mutant in rad52Δ or hta2Δ linkage group
C
True Positive Rate
False Positive Rate
D
% Foci Detected by Classification
% Foci Detected Manually
E
All Screens
yku80Δ - ts
sgs1Δ - ts
SM+Phleo - ts
SM - ts
yku80Δ
sgs1Δ
SM+Phleo
SM
FNR
FPR
Percentage
F
Plate 1
Plate 2
Plate 3
Wild-type Mean % Foci
rad51Δ Mean % Foci
xrs2Δ Mean % Foci
Cells / Sample

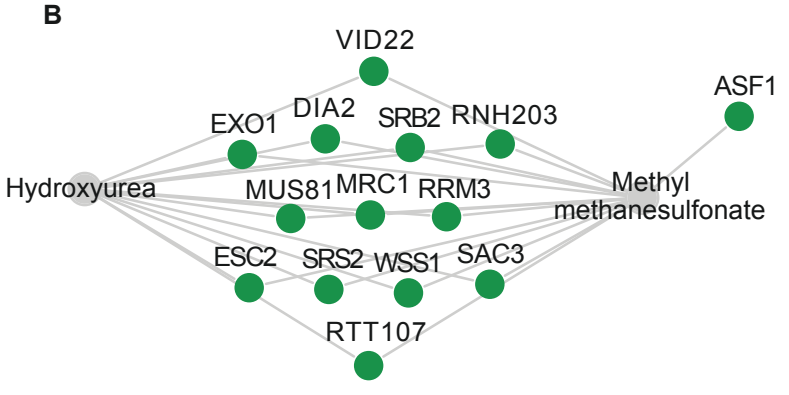
Figure S1. Optimization of Integrated SGA-HCS Pipeline, Related to Figure 1.

A. Growth of a *RAD52- GFP* strain in the presence of hydroxyurea (HU). The tagged query strain (BY4879), an isogenic wild-type (BY4394), and a *rad52*Δ strain (Yeast Deletion Collection) were grown to log phase, serially diluted and spotted onto SC medium (untreated) or SC medium containing 100mM HU.
B. A quantitative fitness assessment of all viable deletion mutants carrying fluorescent markers on the SGA output arrays. Colony sizes of GFP- and RFP-tagged deletion mutants were quantified as a proxy of cellular fitness, and compared to previously scored single mutant fitness values (Baryshnikova et al. 2010). Blue nodes represent tagged single mutant strains that grew comparably to the untagged mutants. Yellow nodes highlight the 100 fluorescently tagged mutants with the most significant fitness defects; green x-nodes indicate genes that are part of the *RAD52* or *HTA2* linkage groups, and red +-nodes indicate strains whose growth defects that can be explained by a known GI with *RAD52* or *HTA2* are indicated with red crosses (4 and 2 mutants, respectively).
C. Receiver operating characteristic (ROC) curve (solid line) illustrating the performance of Rad52 focus classifier compared to random classification (dashed line). Reiterations of training were performed on 1/5th of the training set to generate the ROC curve, identifying an overall cross-validation of 98%.
D. Correlation between automated foci classification and manual foci classification of 50 representative images (Pearson correlation = 0.96).
E. Graph illustrating the performance of the classifier in different screens. The False Negative Rate (FNR; red), and False Positive Rate (FPR; green) for each screen is shown. Examples of cell segmentation errors that produce a False Positive or False Negative are shown. Cell outlines (red and green) are produced by CellProfiler™ segmentation, and represent cells as visualized using CellProfiler Analyst™. The False Positive classification shown reflects erroneous division of a focus into two, while the False Negative classification reflects erroneous division of one cell into two, with the focus isolated in one half.
F. The results of bootstrapping analysis in three mutant backgrounds are shown; Wild-type strain (*his3*Δ; green; top panel), *rad51*Δ (red; middle panel) and *xrs2*Δ (blue; bottom panel). Each mutant is shown in three independent biological replicates (left to right). Error bars indicate the standard deviation in the population average of identified Rad52- GFP foci observed when sampling differently sized cell populations, from a minimum sample size of 10 cells to a maximum sample size of 2000 cells. The dashed line in the center of each graph indicates the sample size of 1000 cells.

A


No known interaction
30%
Physical interaction
7%
Genetic interaction
70%
0
10
20
30
40
Number of Genes
THP2, IST3, SIC1, CLB5, RPL16B, FTR1, CHZ1, NIT1, MRS2, PUS2, SAC3
RAD51, SRS2, RTT107
ELG1, RAD51, VID22, MUS81, MRC1, RNH203, RNH202, SRS2, SRB2, RAD59, RTT107, TSA1, WSS1, SPT2, POL32, XRS2, RTT109, DDC1, RAD17 RRM3, EXO1, RAD55, RAD24, HST3, CTF18, TOF1, UBC4, RAD57, TOP1, MMS1, ASF1, ESC2, DIA2

B


VID22
ASF1
EXO1
DIA2
SRB2
RNH203
Hydroxyurea
MUS81
MRC1
RRM3
Methyl methanesulfonate
ESC2
SRS2
WSS1
SAC3
RTT107

C

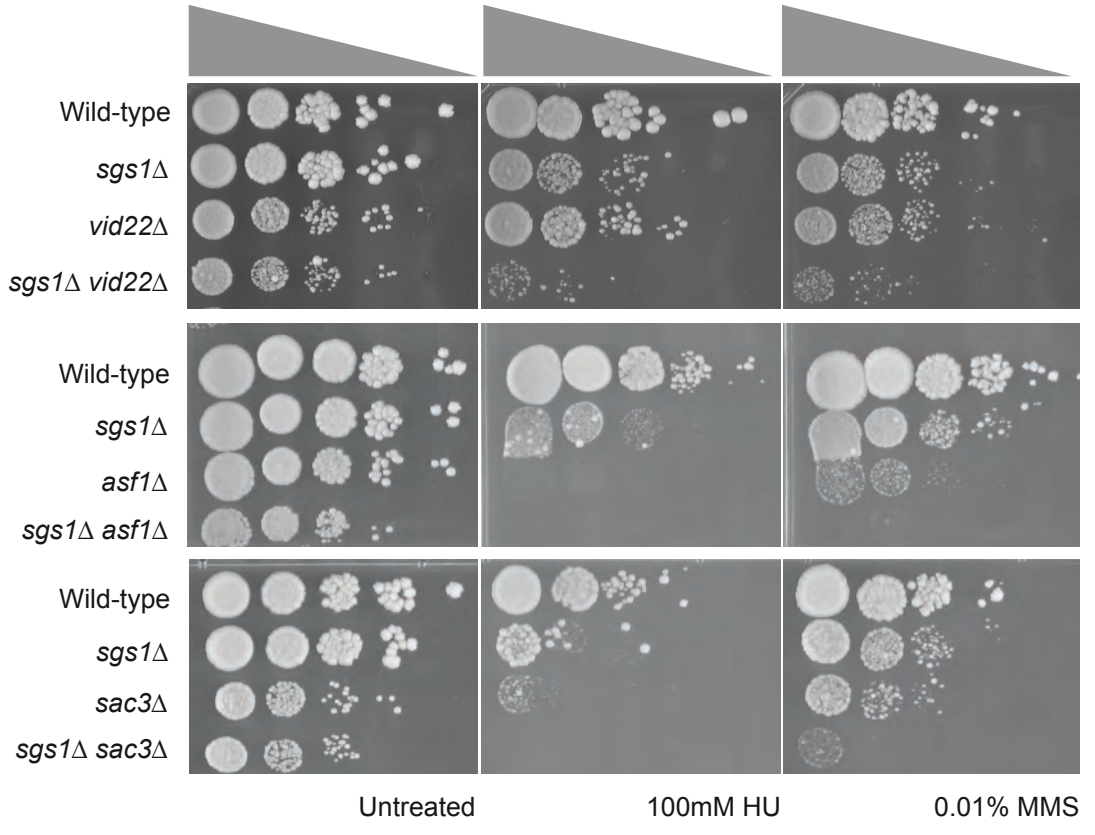

Wild-type
sgs1Δ
vid22Δ
sgs1Δ vid22Δ
Wild-type
sgs1Δ
asf1Δ
sgs1Δ asf1Δ
Wild-type
sgs1Δ
sac3Δ
sgs1Δ sac3Δ
Untreated
100mM HU
0.01% MMS

Figure S2. SGA + HCS Identifies Novel Genetic Interactions with *SGS1*, Related to Figure 2.

A. Summary of the BioGRID-curated genetic and physical interactions between *SGS1* and genes identified in our screens of the non-essential deletion collection for elevated levels of foci. 70% (31/44) of the hits detected physically and/or genetically interact with *SGS1*. Insets list genes that have elevated levels of Rad52-GFP foci as identified in *sgs1*Δ screens that fall into each interaction category.
B. Double mutants identified in *sgs1*Δ double mutant screens of the non-essential deletion collection were grown in the presence of hydroxyurea (HU) or methyl methanesulfonate (MMS). The 14 (of 44 tested) strains that were more sensitive to HU or MMS in an *sgs1*Δ background are summarized in this diagram.
C. The indicated strains were grown to log phase, serially diluted, and spotted onto SC medium (untreated), SC medium +100mM HU, and SC medium +0.01% MMS. Growth of an *sgs1*Δ strain, a *vid22*Δ strain, an *asf1*Δ strain, a *sac3*Δ strain, and all three *sgs1*Δ double mutant strains is shown.

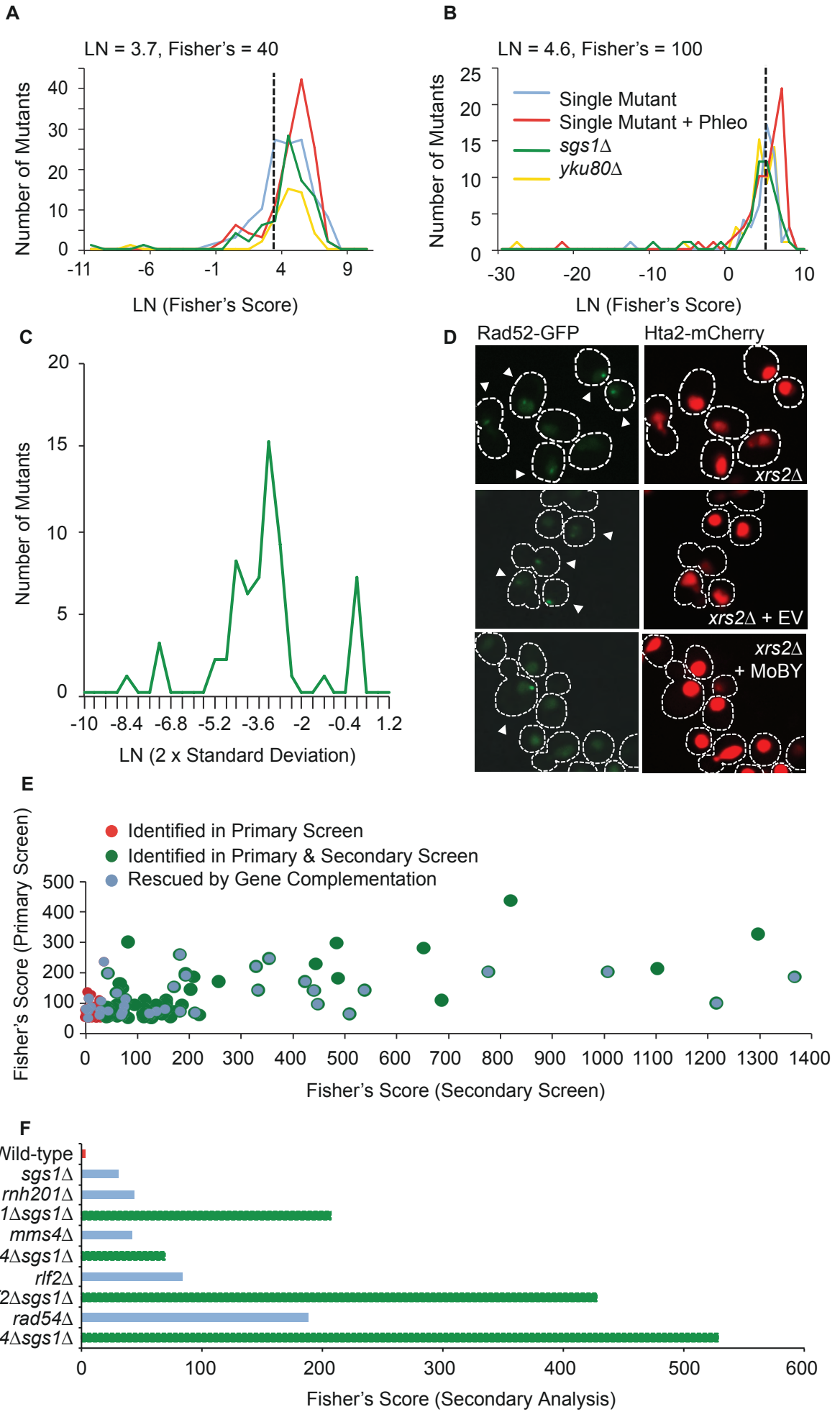
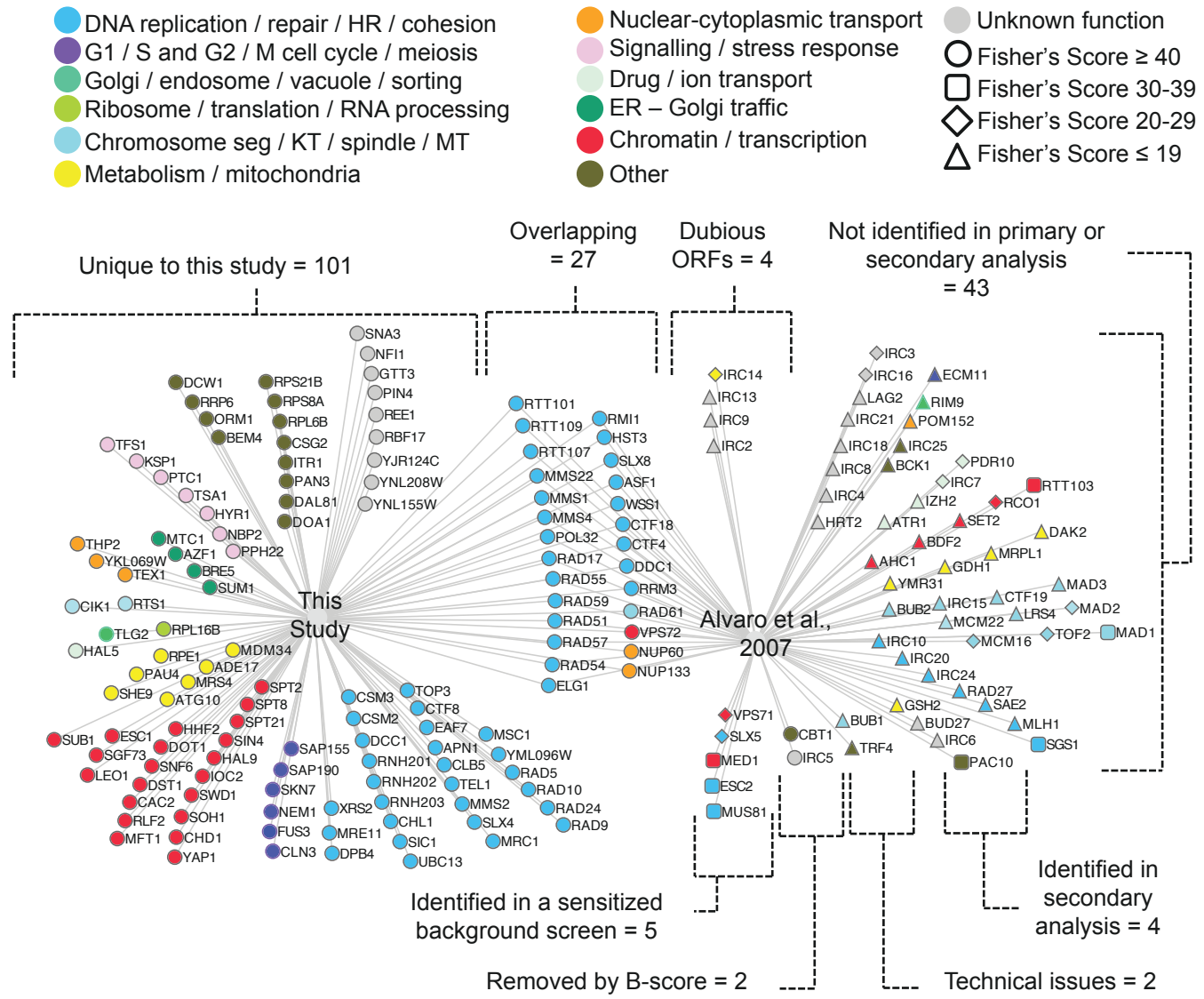

A
LN = 3.7, Fisher's = 40
Number of Mutants
LN (Fisher's Score)
B
LN = 4.6, Fisher's = 100
Single Mutant
Single Mutant + Phleo
sgs1Δ
yku80Δ
C
LN (2 x Standard Deviation)
D
Rad52-GFP
Hta2-mCherry
xrs2Δ
xrs2Δ + EV
xrs2Δ + MoBY
E
Identified in Primary Screen
Identified in Primary & Secondary Screen
Rescued by Gene Complementation
Fisher's Score (Primary Screen)
Fisher's Score (Secondary Screen)
F
Wild-type
sgs1Δ
rnh201Δ
rnh201Δsgs1Δ
mms4Δ
mms4Δsgs1Δ
rlf2Δ
rlf2Δsgs1Δ
rad54Δ
rad54Δsgs1Δ
Fisher's Score (Secondary Analysis)

Figure S3. Secondary Screening and Hit Confirmation via Gene Complementation, Related to Figure 2.

A. The distribution of the scores from each confirmation or mini-array screen using the *natMX*-marked deletion collection (Costanzo et al. 2010) is plotted as a LN function of the Fisher's score, where the cutoff used to define statistically relevant increases in DNA damage foci is LN = 3.7 (dashed line), Fisher's score = 40. 63% of mutants (146/230 tested) maintained a significantly elevated foci phenotype in the secondary, *natMX*-marked deletion collection. Single mutants = blue (BY5084); Single mutants in the presence of phleomycin = red (BY5084); *sgs1*Δ double mutant strains = green (BY5085); *yku80*Δ double mutant strains = yellow (BY5086).
B. The distribution of the scores from each confirmation or mini-array screen using the *natMX*-marked TS-mutant collection is plotted as a LN function of the Fisher's score, where the cutoff used to define statistically relevant increases in DNA damage foci is LN = 4.6 (dashed line), Fisher's score = 100. 66% of mutants (79/119 tested) maintained a significantly elevated foci phenotype in the secondary, *natMX*-marked TS-mutant collection. Single mutants = blue (BY5084); Single mutants in the presence of phleomycin = red (BY5084); *sgs1*Δ double mutant strains = green (BY5085); *yku80*Δ double mutant strains = yellow (BY5086).
C. The change in levels of Rad52-GFP foci as a result of plasmid-based complementation of non-essential single mutants is plotted as a LN function. 46% of mutants (37/80 tested) were rescued by addition of the MoBY-ORF plasmid (constructed in a BY5092 background), using 2x standard deviations from empty vector (EV) control samples / mutant as a cutoff (Ho et al 2009).
D. Images of an *xrs2*Δ mutant (top), an *xrs2*Δ mutant transformed with an empty vector plasmid (EV; middle), and an *xrs2*Δ mutant transformed with the MoBY-XRS2 plasmid (bottom). Green channel images (Rad52-GFP) and red channel images (Hta2-mCherry) are shown, with cells outlined in white dashed lines.
E. Scatter plot mapping of the distribution of Fisher's scores of non-essential single mutant hits. The majority (77/124 tested = 62%) of non-essential single mutant hits were confirmed in at least one secondary assay (red = identified exclusively in primary screen; green = identified in both primary and secondary screen; blue = rescued by gene complementation).
F. Fisher's Score analysis of four reconstructed mutants predicted to be false negatives in the *sgs1*Δ screens (*rnh201*Δ *sgs1*Δ, BY5781; *mms4*Δ *sgs1*Δ, BY5782; *rlf2*Δ *sgs1*Δ, BY5783; *rad54*Δ *sgs1*Δ, BY5784; green) plotted with corresponding single mutants (blue) and a wild-type control strain (red).

A


DNA replication / repair / HR / cohesion
G1 / S and G2 / M cell cycle / meiosis
Golgi / endosome / vacuole / sorting
Ribosome / translation / RNA processing
Chromosome seg / KT / spindle / MT
Metabolism / mitochondria
Nuclear-cytoplasmic transport
Signalling / stress response
Drug / ion transport
ER – Golgi traffic
Chromatin / transcription
Other
Unknown function
Fisher's Score ≥ 40
Fisher's Score 30-39
Fisher's Score 20-29
Fisher's Score ≤ 19
Unique to this study = 101
Overlapping = 27
Dubious ORFs = 4
Not identified in primary or secondary analysis = 43
This Study
Alvaro et al., 2007
Identified in a sensitized background screen = 5
Removed by B-score = 2
Identified in secondary analysis = 4
Technical issues = 2

B

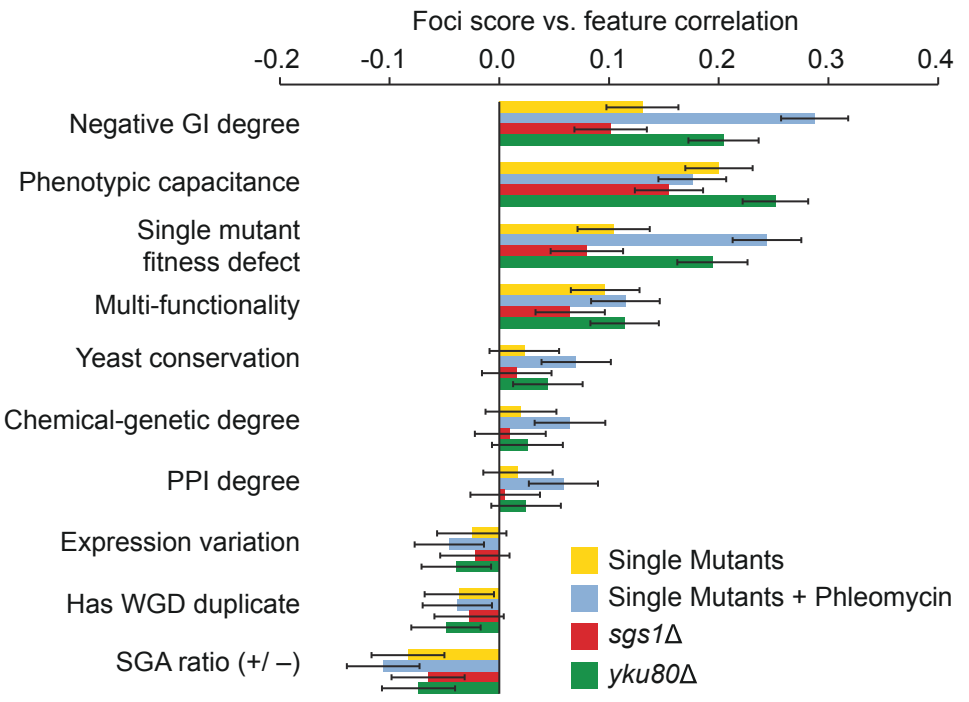
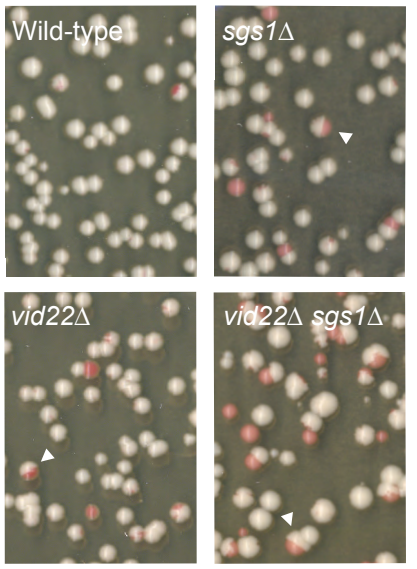

Foci score vs. feature correlation
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
Negative GI degree
Phenotypic capacitance
Single mutant fitness defect
Multi-functionality
Yeast conservation
Chemical-genetic degree
PPI degree
Expression variation
Has WGD duplicate
SGA ratio (+/ –)
Single Mutants
Single Mutants + Phleomycin
sgs1Δ
yku80Δ

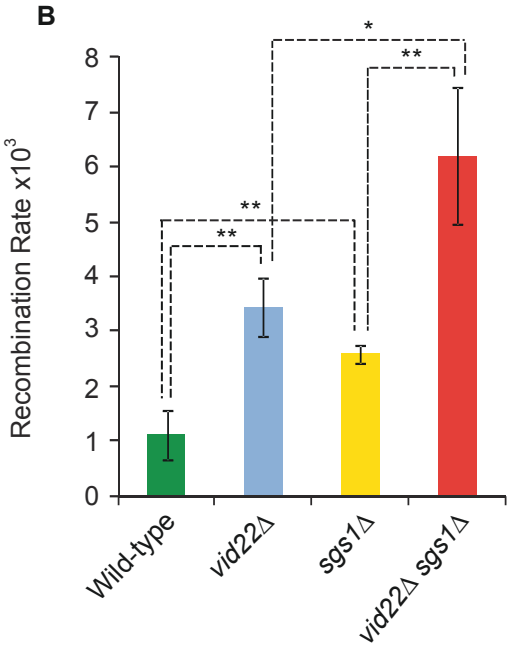
Figure S4: General trends of enrichment and overlap in non-essential data. Related to Figure 3.

A. A direct comparison between the common non-essential single mutants tested in this study and a previous screen that assessed elevated levels of Rad52-YFP foci (Alvaro et al 2007). Mutants identified uniquely in either screen or in both screens are indicated by brackets. Reasons for a failure to detect hits identified in Alvaro et al., 2007 in our screens are highlighted with brackets where applicable. Many hits identified in the parallel screen performed by Alvaro et al were close to the cutoff used in our dataset, but were deemed not statistically reliable enough for inclusion in the high confidence hit list (indicated using non-circular node shapes; shape of node corresponds to a range of Fisher's scores as described in the legend above the network). Nodes are color-coded by biological process (legend at top of figure).
B. Pearson correlation analysis between foci phenotype score (Fisher's Score) and physiological and evolutionary properties of yeast was measured for each screen condition (SM = yellow, SM + Phleo = blue, *sgs1*Δ = red, *yku80*Δ = green).

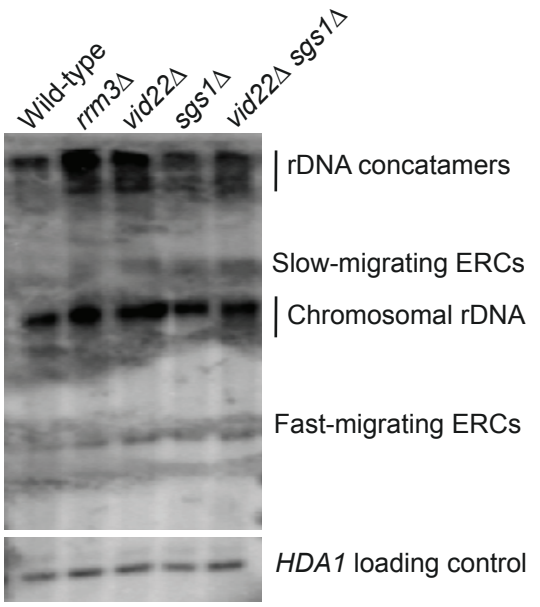
A


Wild-type
sgs1Δ
vid22Δ
vid22Δ sgs1Δ

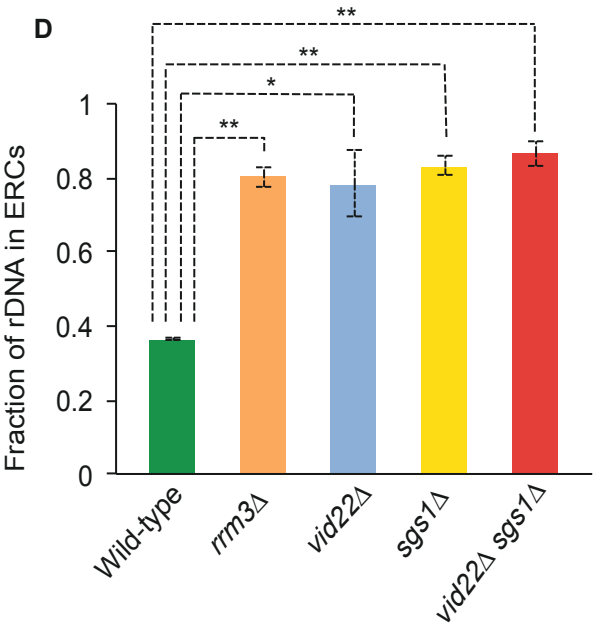
B


Recombination Rate x10^3
0
1
2
3
4
5
6
7
8
*
**
**
**
Wild-type
vid22Δ
sgs1Δ
vid22Δ sgs1Δ

C


Wild-type
rrm3Δ
vid22Δ
sgs1Δ
vid22Δ sgs1Δ
rDNA concatamers
Slow-migrating ERCs
Chromosomal rDNA
Fast-migrating ERCs
HDA1 loading control

D


Fraction of rDNA in ERCs
0
0.2
0.4
0.6
0.8
1
**
**
*
**
Wild-type
rrm3Δ
vid22Δ
sgs1Δ
vid22Δ sgs1Δ

E

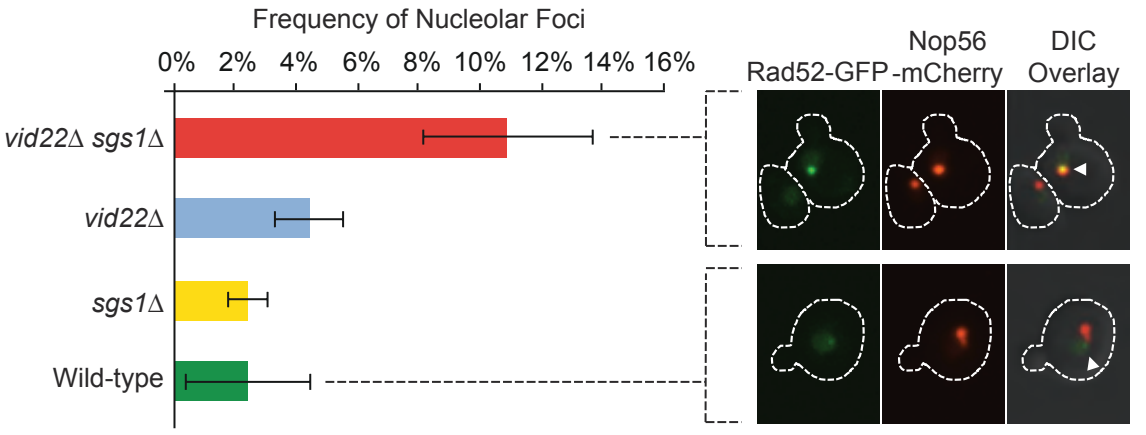
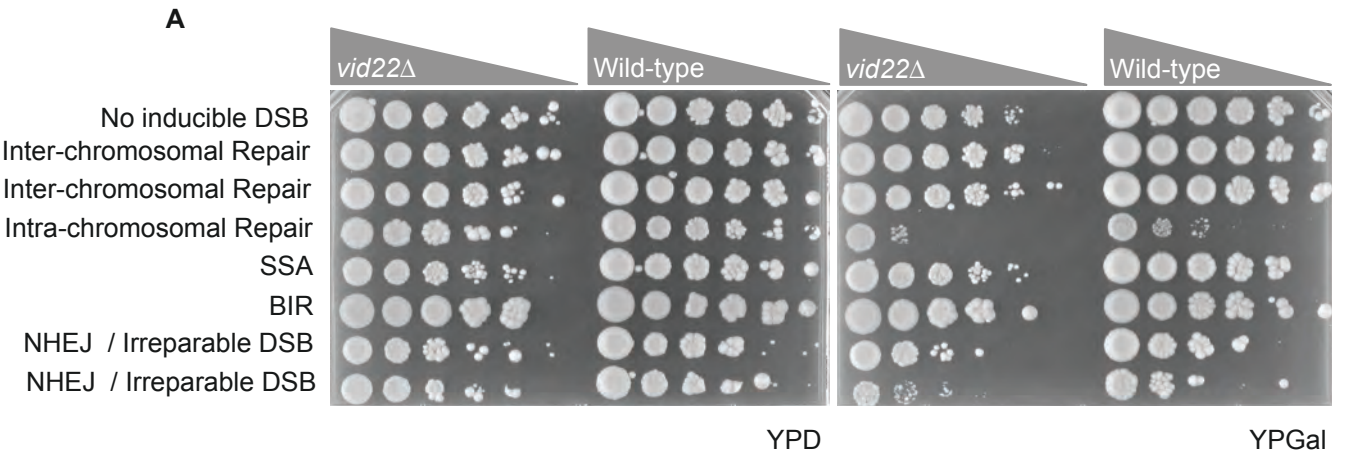

Frequency of Nucleolar Foci
0%
2%
4%
6%
8%
10%
12%
14%
16%
vid22Δ sgs1Δ
vid22Δ
sgs1Δ
Wild-type
Rad52-GFP
Nop56
-mCherry
DIC
Overlay

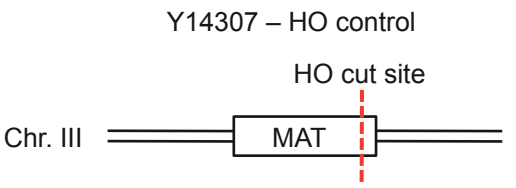
Figure S5. Effect of Deletion of *VID22* and *SGS1* on Stability of the rDNA Locus, Related to Figure 4.

A. Representative images of half sectored colonies in which an *ADE2* mutation was integrated at the rDNA locus in wild-type, *vid22*Δ, *sgs1*Δ, and *vid22*Δ *sgs1*Δ cell populations, as indicated. White arrowheads indicate half sectored colonies from each mutant population.
B. Quantification of *ADE2* marker loss at the rDNA array via colony half sectoring in *vid22*Δ (BY5482), *sgs1*Δ (BY5483), and *vid22*Δ *sgs1*Δ (BY5484) cell populations as compared to wild-type cells (BY5481), as well as in *vid22*Δ *sgs1*Δ cell populations as compared to either *vid22*Δ or *sgs1*Δ cells (wild-type = green, *vid22*Δ = blue, *sgs1*Δ = yellow, *vid22*Δ*sgs1*Δ = red, * = *p-value* <0.05, ** = *p-value* < 0.01, N = 3). Error bars represent the standard deviation between three replicate experiments.
C. Southern blot analysis of ERCs in *Bam*HI digested genomic DNA from wild-type, *rrm3*Δ, *vid22*Δ, *sgs1*Δ and *vid22*Δ *sgs1*Δ cell populations (*rrm3*Δ genomic DNA = positive control, *HDA1* = loading control). The blot was probed with an *Eco*RI digested probe for a single rDNA repeat (plasmid 2484), and the bands representing various forms of rDNA are indicated to the right.
D. Quantification of percentage of total genomic rDNA present as extra-chromosomal circles (ERCs) identified via southern blot analysis. *vid22*Δ, *sgs1*Δ, and *vid22*Δ *sgs1*Δ mutants as compared to wild-type cells (*rrm3*Δ genomic DNA = positive control; *HDA1* = loading control, wild-type = green, BY5479; *rrm3*Δ = orange, Yeast Deletion Collection; *vid22*Δ = blue, BY5480; *sgs1*Δ = yellow, BY5401; *vid22*Δ*sgs1*Δ = red, BY5160; * = *p-value* <0.005, ** = *p-value* < 0.0005, N = 3). Error bars indicate standard deviation in percentage of total genomic rDNA present as ERCs across three biological replicates.
E. Quantification of frequency of sub-nucleolar foci in wild-type (BY5440), *vid22*Δ (BY5442), *sgs1*Δ (BY5441), and *vid22*Δ*sgs1*Δ (BY5443) mutant populations (wild-type = green, *vid22*Δ = blue, *sgs1*Δ = yellow, *vid22*Δ*sgs1*Δ = red, N = 3). Inset: Images of cells showing overlap of Rad52-GFP and Nop56-mCherry (a nucleolar protein) in a *vid22*Δ *sgs1*Δ cell (top) and wild-type cell (bottom). Rad52-GFP foci are indicated with a white arrowhead.

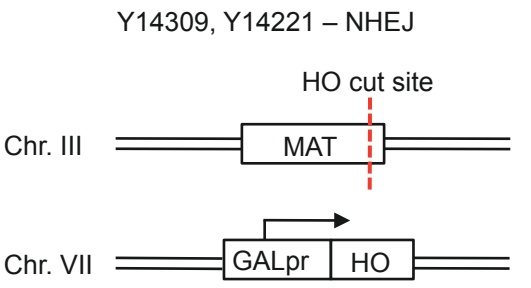
A


vid22Δ
Wild-type
vid22Δ
Wild-type
No inducible DSB
Inter-chromosomal Repair
Inter-chromosomal Repair
Intra-chromosomal Repair
SSA
BIR
NHEJ / Irreparable DSB
NHEJ / Irreparable DSB
YPD
YPGal

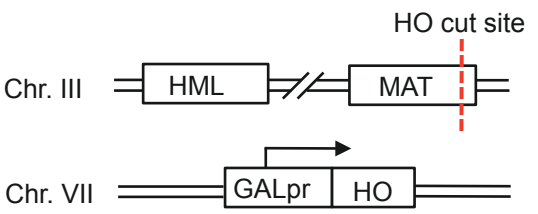
B


Y14307 – HO control
HO cut site
Chr. III
MAT

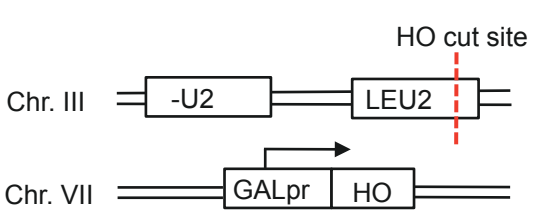
C


Y14309, Y14221 – NHEJ
HO cut site
Chr. III
MAT
Chr. VII
GALpr
HO

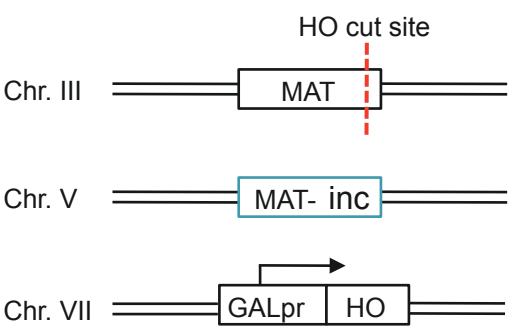
D


Y14220 – HR using
intra-chromosomal recombination
HO cut site
Chr. III
HML
MAT
Chr. VII
GALpr
HO

E


Y14223 – SSA
HO cut site
Chr. III
-U2
LEU2
Chr. VII
GALpr
HO

F


Y14308, Y14222 – HR using
inter-chromosomal recombination
HO cut site
Chr. III
MAT
Chr. V
MAT- inc
Chr. VII
GALpr
HO

G

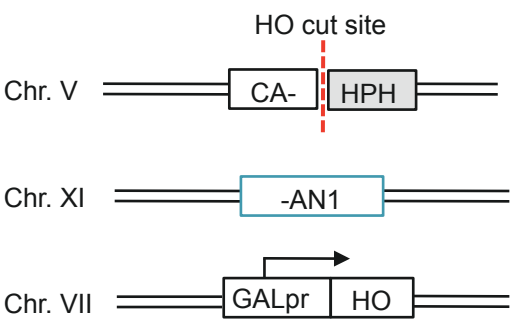

Y14310 – BIR
HO cut site
Chr. V
CA-
HPH
Chr. XI
-AN1
Chr. VII
GALpr
HO

Figure S6. Deletion of *VID22* Leads to Defects in NHEJ, Related to Figure 4.

A. Serial spot dilutions of *vid22*Δ or wild-type versions of a series of strains designed to test different aspects of the DDR are shown (described in Haber at al 2002). The left panel shows growth after plating on YPD medium, and the right panel shows growth after plating on YPGal medium to induce a DNA double-strand break.

B-F) Diagrams illustrating relevant features of strains to query the DDR:

B. Control strain: contains an HO cut site integrated at the 3' end of the MAT locus but does not carry a galactose-inducible *HO* endonuclease (Y14307, Y14318).
C. NHEJ-specific strain: contains an HO cut site integrated at the 3' end of the MAT locus and a galactose-inducible *HO* endonuclease integrated on chromosome seven. Repair via recombination is abrogated due to the deletion of homologous *HML* and *HMR* loci (Y14309, Y14221, Y14330, Y14346).
D. Intra-chromosomal recombination-specific strain: contains an HO cut site integrated at the 3' end of the MAT locus and a galactose- inducible *HO* endonuclease integrated on chromosome seven. The *HMR* locus has been deleted, but the *HML* locus remains and provides a template for recombination (Y14220, Y14327).
E. SSA-specific strain: contains an HO cut site integrated in the middle of the *LEU2* locus on chromosome three and a galactose-inducible *HO* endonuclease integrated on chromosome seven. A 1.3kb fragment containing the 3' end of the *LEU2* locus has been integrated 30kb distal to the endogenous locus, and is represented as –U2 (Y14223, Y14334).
F. Inter-chromosomal recombination-specific strain: contains an HO cut site integrated at the 3' end of the MAT locus and a galactose-inducible *HO* endonuclease integrated on chromosome seven. Although both the *HML* and *HMR* loci have been deleted, a secondary *MAT* locus with a disrupted HO cut site that cannot be cleaved has been integrated on chromosome five, providing a template for repair (Y14308, Y14222, Y14332, Y14338).
G. BIR-specific strain: contains an HO cut site integrated in the middle of the *CAN1* locus on chromosome five, and the sequence proximal to the cut site has been replaced with a hygromycin B phosphotransferase (HPH) resistance cassette; the remaining sequence is represented as CA–. The latter portion of the *CAN1* sequence, which shares 1157-bp homology and is represented as –AN1, has been integrated on chromosome eleven. A galactose-inducible *HO* endonuclease is integrated on chromosome seven. Strain numbers for control strains in the wild-type background are given (Y14310, Y14380).

Table S1. Strains Used in This Study, Related to Experimental Procedures.

Strain ID / Source	Mating Type	Genotype
BY4394	alpha	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ*
Yeast Deletion Collection	**a**	*rad52 Δ::kanMX his3 Δ1 leu2 Δ0 ura3 Δ0 met15Δ0*
BY4879	alpha	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1pr::RPL39pr-tdTomato::CaURA3 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX*
BY4880	alpha	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX sgs1 Δ::RPL39pr-tdTomato::CaURA3*
BY4881	alpha	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX yku80 Δ::RPL39pr-tdTomato::CaURA3*
BY5084	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 met15 Δ0 RAD52-GFP-kanMX HTA2-mCherry-LEU2*
BY5085	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 met15 Δ0 RAD52-GFP-kanMX HTA2-mCherry-LEU2 sgs1 Δ::RPL39pr-tdTomato::CaURA3*
BY5086	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 met15 Δ0 RAD52-GFP-kanMX HTA2-mCherry-LEU2 yku80 Δ::RPL39pr-tdTomato::CaURA3*
BY5092	alpha	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX*
BY5781	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX sgs1 Δ::RPL39pr-tdTomato::CaURA3 rnh201 Δ::kanMX*
BY5782	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX sgs1 Δ::RPL39pr-tdTomato::CaURA3 mms4 Δ::kanMX*
BY5783	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX sgs1 Δ::RPL39pr-tdTomato::CaURA3 rlf2 Δ::kanMX*
BY5784	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX sgs1 Δ::RPL39pr-tdTomato::CaURA3 rad54 Δ::kanMX*
BY5418	alpha	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 CAN1pr::RPL39pr-tdTomato::CaURA3::can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX vid22 Δ::HPH*
BY5433	**a**	*his3 Δ1 leu2 Δ0 ura3 Δ0 MET15 CAN1pr::RPL39pr-tdTomato::CaURA3::can1 Δ::STE2pr-LEU2 lyp1 Δ RAD52-GFP-HIS3 HTA2-mCherry-natMX vid22 Δ::HPH sgs1 Δ:kanMX*

BY5480	**a**	*vid22* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 MET15*
BY4741	**a**	*his3* Δ*1 leu2* Δ*0 ura3* Δ*0 MET15*
BY5401	**a**	*sgs1* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 MET15*
BY5160	**a**	*vid22* Δ::*natMX sgs1* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15* Δ*0*
BY5440	alpha	*NOP56-mCherry-natMX RAD52-GFP-HIS3 his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15* Δ*0*
BY5441	alpha	*NOP56-mCherry-natMX RAD52-GFP-HIS3 sgs1* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15* Δ*0*
BY5442	alpha	*NOP56-mCherry-natMX RAD52-GFP-HIS3 vid22* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15* Δ*0*
BY5443	alpha	*NOP56-mCherry-natMX RAD52-GFP-HIS3 vid22* Δ::*kanMX sgs1* Δ::*HPH his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15* Δ*0*
BY5481	**a**	*RAD5+ RDN1::ADE2 leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100*
BY5482	**a**	*vid22* Δ::*kanMX RAD5+ RDN1::ADE2 leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100*
BY5483	**a**	*sgs1* Δ::*HIS3 RAD5+ RDN1::ADE2 leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100*
BY5484	**a** / alpha	*RAD5+ RDN1/RDN1::ADE2 VID22/vid22* Δ::*kanMX SGS1/sgs1* Δ::*HIS3 leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100*
BY5479	**a**	*his3* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 MET15*
Yeast Deletion Collection	**a**	*rrm3* Δ::*kanMX his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15Δ0*
BY5486	**a**	*his3* Δ*1 leu2* Δ*0 ura3* Δ*0 met15* Δ*0 sir4* Δ*::kanMX trp1* Δ*::HPH*
BY5487	**a** / alpha	*his3* Δ*1/his3Δ1 leu2* Δ*0/leu2Δ0 ura3* Δ*0/ura3Δ0 met15* Δ*0/MET15 LYS2/lys2* ∆*0 sir4* Δ*::kanMX trp1Δ::HPH/trp1Δ::HPH VID22-SIR4::natMX*
BY5507	**a**	*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 rad5-535, TBF1-13xMYC::HIS3* ***W303**
BY5493	**a**	*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 rad5-535,VID22-13xMYC::HIS3* ***W303**
BY5494	**a**	*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 rad5-535, ENV11-13xMYC::HIS3* ***W303**
BY5508	**a**Δ0	*mat* aΔ*::loxP ade2::loxP leu2::pG-HO-L2 lys2::loxP mnt2::LYS2be* ***W303**
BY5495	**a**Δ0	*mat* aΔ*::loxP ade2::loxP leu2::pG-HO-L2 lys2::loxP mnt2::plM63 VID22-13xMYC::HIS3* ***W303**
BY5496	**a**Δ0	*mat* aΔ*::loxP ade2::loxP leu2::pG-HO-L2 lys2::loxP mnt2::plM63 VID22-13xMYC::HIS3 sgs1* Δ*::HPH* ***W303**
BY5498	**a**Δ0	*mat* aΔ*::loxP ade2::loxP leu2::pG-HO-L2 lys2::loxP mnt2::plM63 RNH201-13xMYC::HIS3* ***W303**

BY5501	**a**Δ0	*mat* aΔ*::loxP ade2::loxP leu2::pG-HO-L2 lys2::loxP mnt2::plM63 RNH202-13xMYC::HIS3* ***W303**
BY5504	**a**Δ0	*mat* aΔ*::loxP ade2::loxP leu2::pG-HO-L2 lys2::loxP mnt2::plM63 RNH203-13xMYC::HIS3* ***W303**
Y14220	**a**	*HML* **a** *MAT* **a** *ho* Δ *hmr* Δ*::ADE1 ade1-100 leu2,3-112 lys5 trp::hisG ura3-52 ade3::GAL-HO* ***W303**
Y14221	**a**	*MAT* **a** *hml* Δ*::ADE1 hmr* Δ*::ADE1 ade1-100 leu2,3-112 lys5 trp::hisG ura3-52 ade3::GAL-HO* ***W303**
Y14222	**a**	*MAT* **a** *ho* Δ *hml* Δ*::ADE1 hmr* Δ*::ADE1 arg5,6* Δ*::HPH::MATa-inc ade1-100 leu2,3-112 lys5 trp::hisG ura3-52 ade3::GAL-HO* ***W303**
Y14223	**a**	*MAT* **a**Δ*::hisG ho* Δ *hml* Δ*::ADE1 hmr* Δ*::ADE1 his4::URA3-leu2(Xho1-Kpn1 fragment)-pBR322 leu2::HOcs ade1-100 lys5 ura3-52 ade3::GAL-HO* ***W303**
Y14307	alpha	*hml::ADE1 hmr::ADE1 ade1 leu2-3,112 lys5 trp1::hisG ura3-52* ***W303**
Y14308	alpha	*hml::ADE1 hmr::ADE1 ade3::GAL-HO MATinc-UR3A ade1 leu2-3,112 lys5 trp1::hisG ura3-52* ***W303**
Y14309	alpha	*hml* Δ*::ADE1 hmr* Δ*::ADE1 ade3::GAL-HO ade1 leu2-3,112 lys5 trp1::hisG ura3-52* ***W303**
Y14310	**a**	*mat* a*::HOcs* Δ*::hisG ura3* Δ*851 trp1* Δ*63 sup53* Δ*::leu2* Δ*::kanMX hml* Δ*::hisG hmr* Δ*::ADE3 ade3::GAL10::HO can1,1-1446::HOcs::HPH::* Δ *AVT2 ykl215c::leu2::hisG::can1* Δ*1-289* ***W303**
Y14318	**a**	*vid22* Δ*::kanMX hml::ADE1 hmr::ADE1 ade1 leu2-3,112 lys5 trp1::hisG ura3-52* ***W303**
Y14330	**a**	*MAT* **a** *vid22* Δ*::kanMX hml* Δ*::ADE1 hmr* Δ*::ADE1 ade1-100 leu2-3,112 lys5 trp1::hisG ura3-52 ade3::GAL-HO* ***W303**
Y14332	**a**	*MAT* **a** *vid22* Δ*::kanMX ho* Δ *hml* Δ*::ADE1 hmr* Δ*::ADE1 arg5,6* Δ*::HPH::MATa-inc ade1-100 leu2,3-112 lys5 trp::hisG ura3-52 ade3::GAL-HO* ***W303**
Y14334	**a**	*vid22* Δ*::kanMX MAT* **a**Δ*::hisG ho* Δ *hml* Δ*::ADE1 hmr* Δ*::ADE1 his4::URA3-leu2(Xho1-Kpn1 fragment)-pBR322 leu2::HOcs ade1-100 lys5 ura3-52 ade3::GAL-HO* ***W303**
Y14327	**a**	*vid22* Δ*::kanMX HML* **a** *MAT* **a** *ho* Δ *hmr* Δ*::ADE1 ade1-100 leu2,3-112 lys5 trp::hisG ura3-52 ade3::GAL-HO* ***W303**
Y14338	alpha	*vid22* Δ*::kanMX hml::ADE1 hmr::ADE1 ade3::GAL-HO MATinc-URA3 ade1 leu2-3,112 lys5 trp1::hisG ura3-52* ***W303**
Y14346	**a**	*vid22* Δ*::kanMX hml::ADE1 hmr::ADE1 ade3::GAL-HO ade1 leu2-3,112 lys5 trp1::hisG ura3-52* ***W303**
Y14380	**a**	*mat* a*::HOcs* Δ*::hisG ura3* Δ*851 trp1* Δ*63 sup53* Δ*::leu2* Δ*::kanMX hml* Δ*::hisG hmr* Δ*::ADE3 ade3::GAL10::HO can1,1-1446::HOcs::HPH::* Δ *AVT2 ykl215c::leu2::hisG::can1* Δ*1-289 vid22* Δ*::natMX* ***W303**

Table S2 is available as an Excel file.

Table S3. Calculation of False Negative Rate by Analysis of Protein Complexes Divided among Different Screens, Related to Figure 2.

Protein complexes as curated by (2).

Legend:

✖?	"Missed" / false negative complex member (N = 69)
	Complex member counted towards total (N = 254)
	Discarded, uninformative data (N = 94)

Complex	Screening Condition			
	Single Mutant	Phleomycin	*sgs1Δ*	*yku80Δ*
Cdc28/Clb5 complex	CLB5	CLB5	CLB5	CLB5
Cdc73/Paf1 complex	LEO1	LEO1		
Chromatin assembly complex	RLF2 CAC2 ASF1 ✖ ✖	RLF2 ✖ ASF1 HTZ1 ✖	✖ ✖ ASF1 ✖ CHZ1	RLF2 CAC2 ✖ ✖ ✖
COMPASS complex	✖ SWD1			BRE2 ✖
Cytoplasmic ribosomal large subunit	RPL6B RPL16B	RPL6B ✖	✖ RPL16B	
Cytoplasmic ribosomal small subunit	RPS8A RPS21B ✖	RPS8A ✖ RPS26B		
Cytoskeleton assembly pathway	BEM4	BEM4		BEM4
DNA polymerase delta complex	POL32	POL32	POL32	POL32
DNA replication factor C complex (Ctf18/Ctf8/Dcc1)	CTF18 CTF8 DCC1 ELG1 RAD24	CTF18 CTF8 DCC1 ELG1 RAD24	CTF18 ✖ ✖ ELG1 RAD24	CTF18 CTF8 DCC1 ELG1 RAD24
Holliday junction resolvase complex	MMS4 ✖		✖ MUS81	
Hst3-Hst4	HST3		HST3	HST3
	SIN4	✖	✖	✖

Kornberg's mediator (SRB) complex	✖ ✖ SOH1	✖ MED1 ✖	SRB2 ✖ ✖	✖ MED1 ✖
Mms22-Rtt101-Mms1	MMS1 RTT101 MMS22	MMS1 RTT101 MMS22	MMS1 ✖ ✖	MMS1 RTT101 MMS22
MRX complex	XRS2 MRE11	XRS2 MRE11	XRS2 ✖	XRS2 ✖
NuA4 histone acetyltransferase complex	EAF7	EAF7		
Nuclear cohesin complex	RAD61	RAD61		
Nuclear pore	NUP133 NUP60 ✖		✖ ✖ SAC3	NUP133 ✖ SAC3
Nucleosomal protein complex	HHF2 ✖ ✖	✖ HHF1 HHT1		
Nucleotide-excision repair factor 1 complex	RAD10	RAD10		RAD10
Rad17/Ddc1/Mec3 complex	DDC1 RAD17	DDC1 RAD17	DDC1 RAD17	DDC1 RAD17
Rad5-Rad6-Rad18	RAD5	RAD5		RAD5
Rad51-Rad57	RAD51 RAD55 RAD54 RAD57	RAD51 RAD55 RAD54 RAD57	RAD51 RAD55 ✖ RAD57	RAD51 RAD55 RAD54 RAD57
Rad9-Chk1	RAD9			RAD9
RecQ helicase-Topo III complex	TOP3 NCE4			TOP3 ✖
Ribonuclease H2 complex	RNH203 RNH202 RNH201		RNH203 RNH202 ✖	RNH203 ✖ ✖
Rtt109-Asf1	RTT109	RTT109	RTT109	RTT109
SAGA complex	CHD1 SGF73	CHD1 ✖		
Slx1/Slx4 complex	SLX4			SLX4
Slx5/Slx8 complex	SLX8 ✖			SLX8 HEX3
SWI/SNF complex	SNF6			✖

Swi/Snf complex	✖			SNF5
Swr1 complex	VPS72 ✖	VPS72 VPS71		
THO complex	THP2 MFT1	THP2 ✖	THP2 ✖	THP2 MFT1
Tof1/Mrc1/Csm3 complex	MRC1 ✖ CSM3	MRC1 ✖ CSM3	MRC1 TOF1 ✖	MRC1 TOF1 CSM3
Ubiquitin conjugating enzyme complex	UBC13 MMS2	UBC13 MMS2		UBC13 ✖
Ubp3/Bre5 complex	BRE5	BRE5		
Carboxy-terminal domain protein kinase complex		CTK1		CTK1
Golgi transport complex		COG6 COG5		COG6 ✖
HOPS complex		VAM6 ✖		✖ VPS41
Lge1/Bre1 complex		LGE1 ✖		✖ BRE1
Rpd3L complex		SDS3 ✖ ✖		✖ SIN3 SAP30
Set3 complex		✖ SET3 SNT1		HST1 ✖ ✖
SCF ubiquitin ligase complex			DIA2	DIA2

Table S4 is available as an Excel file.

Table S5 is available as an Excel file.

Supplemental Experimental Procedures

Strain Construction and Confirmation that Tags did not Impair Protein Function. To visualize DNA damage foci within the cell, we fused a *GFP::HIS3* cassette to the C-terminus of the endogenous *RAD52* gene using polymerase chain reaction (PCR; denaturation: 94°C, 30 seconds; annealing: 52°C, 30 seconds; elongation: 68°C, 4min X 30 cycles) and standard methods for yeast transformation (Gietz and Woods, 2002). To test for a possible growth defect associated with the fluorescent tag, a saturated Rad52-GFP culture was serially diluted 10-fold five times, spotted on synthetic complete (SC) media and SC media containing 100mM hydroxyurea (HU), and growth was assessed after 2 and 3 days respectively (Figure S1A; strains BY4394, BY4879, and a *rad52*Δ strain from the Yeast Deletion Collection; Giaever et al., 2002). To provide spatial and cell cycle context, we fused an mCherry::*NatMX* cassette to the endogenous *HTA2* locus to mark the nucleus and *RPL39*pr-tdtomato::*CaURA3* (Figure 1A; Chong et al., 2015) was integrated into the *CAN1* promoter locus to mark the cytoplasm. Fitness of cells in SGA output arrays was also assessed, by quantifying colony sizes using customized software (Qt Colony Imager V20090716) and comparing these values to single mutant fitness scores (Figure S1B; Costanzo et al., 2010). Of 117 previously identified synthetic lethal interactions with *RAD52* and 34 synthetic lethal interactions with *HTA2*, 113 and 31 mutants had no growth defect when combined with the *RAD52-GFP, HTA2-RFP* tagged strain, including our positive control, *rad51*Δ (BY4879). To generate query strains containing deletions of *SGS1* or *YKU80*, the *RPL39*pr-*RFP::CaURA3* cassette was integrated into the *SGS1* or *YKU80* locus instead of the *CAN1* promoter locus (BY4880, BY4881). To provide SGA compatibility, all fluorescent reporters were generated in a BY4741 SGA query strain background (BY4394; *MAT*α *his3*Δ*1 leu2*Δ*0 ura3*Δ*0 MET15 can1*Δ*::STE2*pr-*LEU2 lyp1*Δ).

Scoring Enrichment and Underrepresentation of Non-essential Mutants: Hits were scored for significant enrichment and underrepresentation by inputting hit lists into FuncAssociate 2.0: The Gene Set Functionator (Berriz et al., 2003), available at http://llama.mshri.on.ca/funcassociate/. LOG Odds (LOD) ratios were calculated by comparison to a manually generated associations file, using the algorithm for an unordered gene list, and calculating both under- and over-enrichment. One thousand simulations were performed, and a significance cutoff of 1 was employed to identify scores and P-values for all input categories. Enrichment of our data was calculated within the functional categories specified in Costanzo et al., 2010, and in other pre-existing DNA damage screens (Figure 3A and Table S4; Alvaro et al., 2007; Aouida et al., 2004; Askree et al., 2004; Begley et al., 2002; Bennett et al., 2001; Chang et al., 2011; 2002; Gatbonton et al., 2006; Hartman and Tippery, 2004; Hillenmeyer et al., 2008; Huang et al., 2002; Levy and Siegal, 2008; Parsons et al., 2004; Stirling et al., 2011; Woolstencroft et al., 2006; Yuen et al., 2007).

Pearson Correlation of Hits with Phenotypic and Evolutionary Traits. The relationship between foci score (Fisher's score) and several gene / protein-level features was computed to characterize the properties of genes implicated in the DNA damage response pathway. For each quantitative feature described below, the Pearson correlation coefficient (PCC) between the foci score and the 3885 array genes was calculated (Figure S4B).

- Negative genetic interaction (GI) degree: negative interactions were used directly from published SGA data (Costanzo et al., 2010).
- Phenotypic capacitance: Used directly from (Levy and Siegal, 2008), and summarizes variance across a range of morphological phenotypes upon deletion of each non-essential gene.
- Single mutant fitness defect: Single mutant fitness for all non-essential deletion mutants was derived from mutant colony size data as described (Baryshnikova et al., 2010; Costanzo et al., 2010). The fitness defect (1-f_i) for a single mutant fitness (f_i) was used.
- Multi-functionality: A quantitative standard for gene multi-functionality was defined from annotations to "biological process" terms of the Gene Ontology. The total number of annotations across the set of functionally distinct GO terms was used as a multi-functionality index (Costanzo et al., 2010; Myers et al., 2006).
- Yeast conservation: the number of species that possess an ortholog of a given gene, when considering 23 divergent species of Ascomycota fungi (measure described with the term "persistence"), and the corresponding ortholog data were downloaded from www.broadinstitute.org/regev/orthogroups/. The 23 species are an expanded set of the original 17 species described previously (Wapinski et al., 2007), with the additions of *S. octosporus, S. japonicus, L. elongosporus, C. parasilosis, C. tropicalis* and *C. guilliermondii.*

- Chemical-genetic degree: data measuring the sensitivity of all non-essential deletion mutants to a library of drugs, and a variety of environmental conditions were used (Hillenmeyer et al., 2008). The number of drug and environmental sensitivities for a specific deletion mutant in the homozygous dataset that met a minimum cutoff of *P-value* <0.05 were summed.
- Protein-protein interaction degree (PPI) is the number of physical interactions reported in BioGRID, version 2.0.58 (Stark et al., 2006) and consists of: Affinity Capture-MS, Affinity Capture-RNA, Affinity Capture-Western, Biochemical Activity, Co-crystal Structure, Co-fractionation, Co-localization, Co-purification, Far Western, FRET, PCA, Protein-peptide, Protein-RNA, Reconstituted Complex, and Two-hybrid.
- Expression variation: represents the average number of mRNA copies of each transcript per cell as assessed in (Holstege et al., 1998).
- Whole genome duplicate (WGD): the list of duplicate pairs is comprised of those identified as the result of a whole genome duplication event (Byrne and Wolfe, 2005). Additionally, any pair of genes fulfilling established similarity requirements (Gu et al., 2002) was also considered a duplicate pair resulting from a small scale duplication event. Specifically, a gene pair must have sufficient sequence similarity score (FASTA Blast, E = 10), and sufficient protein alignment length (>80% of the longer protein). A pair must also have an amino acid level identity of at least 30% for proteins with aligned regions longer than 150 a.a., and $0.01n + 4.8L - 0.32^{(1+\exp(-L/1000))}$ for shorter proteins, where L is the aligned length, and $n = 6$ (Gu et al., 2002; Rost, 1999). Pairs from the WGD event were combined with pairs determined through sequence alone.
- SGA Ratio: a measure of LOG(positive interactions / negative interactions) for each non-essential mutant (Costanzo et al., 2010).

Secondary Analysis of Mutants in the *NatMX*-marked Mutant Collection. All mutants identified in primary screens with increased levels of DNA damage-induced foci were compiled into mini-arrays and assessed in a parallel yeast mutant collection marked with the *NatMX* antibiotic resistance cassette. Four mutants that were identified as hits in primary screens could not be assessed in this way, as they were not present in the *NatMX* collection (*GTT3, SNA3, SOH1, PRI1*). Query strains marked with *GFP::KAN*, *RFP::LEU2* and *RFP::CaURA3* (BY5084, BY5085, BY5086) were crossed to a *NatMX*-marked collection of non-essential deletion mutants (Costanzo et al., 2010), as well as a collection of essential TS-mutants (Li et al., 2011), both containing a *MAT***a**-specific promoter driving *SpHis5*. SGA, imaging and analysis were performed as described previously. In order to identify and rank hits in this collection, a statistical analysis identical to that used in the analysis of primary screens was employed.

Confirmation of Single Mutants with Increased Levels of Foci by Gene Complementation. Single mutants identified with increased levels of DNA damage-induced foci were confirmed using a gene complementation assay. Eighty single gene deletion mutants marked with *GFP::HIS3* and *RFP::NatMX* (constructed in a BY5092 background) were transformed with plasmids from the Molecular Barcoded Yeast (MoBY) ORF collection (Ho et al., 2009). The remaining 48 mutants could not be assessed, as the predicted size of the ORF in the MoBY plasmids was not successfully confirmed via plasmid digest. Each deletion mutant of interest was transformed with a MoBY-ORF plasmid as well as an empty vector (EV) control plasmid. Transformants were grown on SD-U plates, and subsequently replica plated onto SD-UH+N medium. Colonies transformed with MoBY-ORF plasmids as well as EVs were imaged in at least quintuplicate. In order to identify the successful rescue of deletion mutants by gene complementation, the standard deviation of the percentage of foci in the population was calculated between a given deletion mutant in the primary screen and that mutant following transformation with the EV plasmid. Mutants were considered to have been rescued if the percentage of foci in the deletion mutant following transformation with the MoBY-ORF plasmid was 2 standard deviations or higher from the percentage of foci in the mutant following transformation with the EV plasmid.

Calculation of False Negative Rate (FNR). FNR was calculated by assessing the division of protein complexes between screens. If any member of a protein complex was identified in any screen (single mutant or sensitized backgrounds), all other screens were assessed for the identification of members of the same protein complex. Any discrepancies in the specific members of the complex identified between screens were labeled as false negatives. If no members of the complex were identified in a given screen background, this was not labeled as a FN hit, but rather discounted from the calculation as uninformative data. FNR was determined by dividing the total number of complex members that were "missed" by the total number of complex members that should have been identified in

all screens, giving a FNR of 27% (Table S3). A selection of four mutants identified as FN in the *sgs1*Δ screens were reconstructed (BY5781-84) to confirm that a Rad52-focus phenotype was present, indicating that these mutants were incorrectly assigned as "negative" in the primary screen analysis (Figure S3F).

Assessment of Chemical Sensitivity of *sgs1*Δ Double Mutants. A mini-array featuring the top non-essential hits identified in the *sgs1*Δ screens was created using standard SGA pinning technology (see above), by crossing in either a wildtype query strain or an *sgs1*Δ query strain, in 3 biological replicates. Final selection arrays were pinned onto either SC media with no drug, SC media + 100mM HU, or SC media + 0.01% methyl methanesulfonate (MMS), and colony size was assessed using Qt Colony Imager V20090716. Colony growth defects were scored and normalized as previously described (Costanzo et al., 2010) and significant interactions were scored by the following calculation:

$$\text{Interaction Score} = \frac{(\text{mean score of } sgs1\Delta genex\Delta + \text{drug} \,/\, \text{mean score of } sgs1\Delta genex\Delta - \text{drug})}{(\text{mean score of } genex\Delta + \text{drug} \,/\, \text{mean score of } genex\Delta - \text{drug})}$$

A threshold interaction score of 0.9 was used, and any double mutants with a score less than this threshold were confirmed by performing serial 10-fold dilutions of saturated cultures, and spotting them onto synthetic complete (SC) media, SC media containing 100mM HU, and SC media containing 0.01% MMS, and assessed for sensitivity after 2, 4, and 9 days, respectively.

Chromatin Immunoprecipitation of Vid22. To assess localization of proteins to Gal-inducible HO breaks, cells were grown in YPLG (lactic acid / glycerol) medium for 3 hours, followed by induction of HO endonuclease expression by addition of galactose to the medium (2%). ChIP to breaks (strains BY5508, BY5495, BY5496, BY5498, BY5501 and BY5504) and ChIP-Seq (strains BY5507, BY5493, BY5494 and BY5508) assays were carried out as previously described (Ribeyre and Shore, 2012). To confirm that HO-mediated cleavage was occurring, southern blots were performed on genomic DNA digested with *Eco*RV (Figure 4B). DNA samples were run on a 0.8% agarose gel, and transferred to Hybond N+ nylon membrane. The blot was probed with both a ^{32}P-radiolabeled *ADE2* DNA fragment and a ^{32}P-radiolabeled *NMD5* fragment. To determine enrichment of promoter regions in the Vid22 ChIP, fold enrichment was calculated after normalization to both the input fraction and an internal control (*SNR52*).

Kinetic Live Cell Imaging. Strains BY4879, BY4880, BY5418 and BY5433 were grown to an optical density at 600 nm of 0.4 in YPD and imaged using a spinning disc confocal system (WaveFX; Quorum) on a Leica DMI 6000B microscope with Velocity 4 software (PerkinElmer). Images were captured at 30 min intervals in microfluidic chambers (CellAsic; Y04C ONIX plates) with constant flow of YPD at room temperature for 8 hrs. Each image represents the projection of eleven 0.4 mm z-stacks in the DIC, GFP and RFP channels. Images were merged, GFP and RFP levels were adjusted to optimize foci visualization, and image sequences were made in ImageJ 1.45s.

Assessment of Sub-nucleolar Rad52-GFP Foci. Wild-type (BY5440), *vid22*Δ (BY5442), *sgs1*Δ (BY5441), and *vid22*Δ *sgs1*Δ (BY5443) strains harboring Rad52-GFP and Nop56-mCherry were grown to mid-log phase in SC +G418 medium and imaged using a spinning disc confocal system (WaveFX; Quorum) on a Leica DMI 6000B microscope with Velocity 4 software (PerkinElmer). A minimum of 850 cells was imaged in each strain background, and foci were quantified manually. The presence of sub-nucleolar foci was assessed via the co-localization of Rad52-GFP and Nop56-mCherry, which localizes to the nucleolus.

rDNA Unequal Sister Chromatid Exchange Assay. Rate of loss of an *ADE2* marker integrated into the rDNA array was used to measure the instability at the rDNA locus (Kaeberlein et al., 1999). Wild-type (BY5481), *vid22*Δ (BY5482), *sgs1*Δ (BY5483), and *vid22*Δ *sgs1*Δ (BY5484) strains were grown overnight and then plated onto solid YPD with 12.5 µg/ml adenine. Colonies were grown 3-4 days at 28°C, and then placed at 4°C for 3 days prior to analysis. The number of half-red/half-white colonies was determined; each was assumed to represent a marker loss event during the first cell division after plating. The number of half-sectored colonies divided by the total number of colonies (excluding entirely red colonies) was reported as the rate of marker loss. About 10,000-15,000 colonies were examined for each strain in each experiment.

Extrachromosomal rDNA Circle Analysis via Southern Blot. Genomic DNA was isolated from strains BY5479, BY5480, BY5401, BY5160 and an *rrm3Δ* strain from the Yeast Deletion Collection (Giaever et al., 2002) and prepared for Southern blot analysis as described (Medvedik and Sinclair, 2007), except for the addition of a phenol-chloroform extraction step and subsequent ethanol precipitation following RNase treatment. At this time, 1µL of glycogen and 1/10th of the volume of 3M sodium acetate was also added, to aid in efficient DNA precipitation in the absence of tRNA. DNA was digested for three hours at 37°C using *Bam*HI (New England BioLabs, #R0136S), which does not cut within the rDNA cassette, and was analyzed in a 0.7% agarose gel (Certified Megabase Agarose, BioRad; #161-3108). DNA was then transferred to Hybond N+ nylon membrane (Amersham; GE Healthcare Life Sciences; #RPN82N). Plasmid 2484 (originally pNL47; Sinclair, 1997) was digested for three hours at 37°C using *Eco*RI (New England BioLabs, #R0101S) and prepared for use as a ^{32}P-radiolabeled probe to the rDNA repeat using the Prime-It® II Random Primer Labeling Kit (Agilent Technologies; #300385) and Spin-*Pure*™ G-50 Columns (Pure Biotech; #SCD50-50). ^{32}P-radiolabeled *HDA1* DNA was used as a loading control.

Sensitivity to Gal-inducible HO Breaks. Strains Y14220-223, 307-310, 318, 327, 330, 332, 334, 338, 346, and 380, which each carry a galactose (GAL)- inducible allele of the homothallic (HO) endonuclease (Figure S6B-G; described in Haber, 2002) were grown overnight in YEPR (yeast extract peptone raffinose). Saturated cultures were serially diluted 10- fold, and spotted onto YEPG (yeast extract peptone galactose) and YEPD. Strain growth was assessed for sensitivity after three days.

Calling Card Analysis of Vid22. Calling card analysis was performed as described (Wang et al., 2011). In brief, Vid22 was tagged with the component of the Sir4 protein that physically interacts with the Ty5 integrase, in a *sir4*Δ background (BY5487, with control BY5486), transposition was induced, and genomic insertions of the transposon were selected. The integration sites of Ty5 transposons were then mapped using paired-end DNA sequencing, and are detailed in full in Table S5. In order to compare directly to Vid22 ChIP-Seq results and DNA elements, all Vid22-Sir4-directed Ty5 integration peaks for which more than one possible gene target was identified were reduced to a single putative hit. The 96 gene promoters chosen for this reduced list were those that were either closest to the Vid22-Sir4-directed Ty5 integration peak, or had a corresponding hit in the Vid22 ChIP-Seq dataset.

Association between Vid22 DNA-binding Sites and DNA Elements. To identify potential biological functions for Vid22 at specific loci, the association of Vid22 with some known genomic features was analyzed. Given the query sets of all possible Vid22 ChIP binding sites and calling card binding sites, four reference sets (G-quadruplex DNA, γH2A sites, loci with elevated basal levels of RNA-DNA hybrids, and loci with elevated levels of RNA-DNA hybrids in an *rnh1*Δ *rnh202*Δ double mutant background) were assessed to identify the number of overlapping regions between the reference and query sets. Regions of G-quadruplex (G4) DNA and γH2A sites in the yeast genome were assessed based on previously reported data (Capra et al., 2010). Direct overlap of ChIP-Seq and calling card binding regions with these genomic structures was assessed using an expanded form of the regions (500 bp up and downstream of the G4 or γH2A site), since these genomic features are very short (average length is 60.9bp ±36.8bp and 57.9bp ±2.9bp, respectively). Loci with elevated levels of RNA-DNA hybrid formation were assessed at the ORF level rather than precise overlapping sequence information, as sequence information was not available for these features (Chan et al., 2014). Fold enrichment was calculated using the following formula:

$$\text{Fold Enrichment} = (s / S) / (p / P)$$

in which *s* represents the number of successes in the given sample (e.g. number of G4 sites that overlap with Vid22 ChIP binding sites), *S* represents the total sample size (e.g. the total number of Vid22 ChIP binding sites), *p* represents the number of successes in the population (e.g. the total number of G4 DNA regions in the genome), and *P* represents the total population size (e.g. the total number of possible G4 regions, based on the cumulative size of the genome). In the case of both G4 regions and γH2A sites, the average size of the feature including the expanded 500bp window was taken into account to identify the total number of possible sites in the genome (i.e. *P* = Total length of the genome / Average length of expanded feature). In the case of loci with elevated levels of RNA-DNA hybrids, the total number of yeast ORFs was used as the total population size (i.e. *P* = 6117).

Genetic Interaction Analysis of *VID22*. Data are available at http://andrewslab.ccbr.utoronto.ca/supplement/styles2015/. Username = styles2015, Password = microscopy.

CellProfiler™ Pipeline

Module #1: LoadImages

- Text-Exact match: Type the text that one type of image has in common (for TEXT options), or their position in each group (for ORDER option): *.flex*
- What do you want to call these images within CellProfiler? *Orig*
- Type the text that one type of image has in common (for TEXT options), or their position □in each group (for ORDER option). Type "Do not use" to ignore: *Do not use*
- What do you want to call these images within CellProfiler? (Type "Do not use" to ignore) □*Do not use*
- Type the text that one type of image has in common (for TEXT options), or their position in each group (for ORDER option): *Do not use*
- What do you want to call these images within CellProfiler? *Do not use*
- Type the text that one type of image has in common (for TEXT options), or their position □in each group (for ORDER option): *Do not use*
- What do you want to call these images within CellProfiler? *Do not use*
- If using ORDER, how many images are there in each group (i.e. each field of view)? *3*
- What type of files are you loading? *tif,tiff,flex movies*
- Analyze all subfolders within the selected folder? *Yes*
- Enter the path name to the folder where the images to be loaded are located. Type period □(.) for default image folder. . □

Module #2: GroupMovieFrames

- What did you call the movie you want to extract from? *Orig*
- How many frames should be extracted each cycle? *2*
- Are the frames grouped by cycle interleaved (ABCABC...) or separated (AA..BB..CC..)? *Interleaved*
- What do you want to call frame 1 in each cycle (or "Do not use" to ignore)? *GFP*
- What do you want to call frame 2 in each cycle (or "Do not use" to ignore)? *RFP*
- What do you want to call frame 3 in each cycle (or "Do not use" to ignore)? *Do not use*
- What do you want to call frame 4 in each cycle (or "Do not use" to ignore)? *Do not use*
- What do you want to call frame 5 in each cycle (or "Do not use" to ignore)? *Do not use*
- What do you want to call frame 6 in each cycle (or "Do not use" to ignore)? *Do not use*

Module #3: RescaleIntensity

- What did you call the image to be rescaled? *RFP*□
- What do you want to call the rescaled image? *RescaledRFP*□
- Rescaling method. (S) Stretch the image (0 to 1). (E) Enter the minimum and maximum values in the boxes below. (G) rescale so all pixels are equal to or Greater than one. (M) Match the maximum of one image to the maximum of another. (C) Convert to 8 bit. (T) Divide by loaded text value. See the help for details. *Stretch 0 to 1*□
- (Method E only): Enter the intensity from the original image that should be set to the lowest value in the rescaled image, or type AA to calculate the lowest intensity automatically from all of the images to be analyzed and AE to calculate the lowest intensity from each image independently. *AA*
- (Method E only): Enter the intensity from the original image that should be set to the highest value in the rescaled image, or type AA to calculate the highest intensity automatically from all of the images to be analyzed and AE to calculate the highest intensity from each image independently. *AA*
- (Method E only): What value should pixels at the low end of the original intensity range be mapped to (range [0,1])? *0*
- (Method E only): What value should pixels at the high end of the original intensity range be mapped to (range [0,1])? *1*□(Method E only): What value should pixels *below* the low end of the original intensity range be mapped to (range [0,1])? *0*
- (Method E only): What value should pixels *above* the high end of the original intensity range be mapped to

(range [0,1])? *1* (Method M only): What did you call image whose maximum you want rescaled image to match? *Orig*

- (Method T only): What did you call the loaded text in the LoadText module?
-

Module #4: RescaleIntensity

- What did you call the image to be rescaled? *GFP*
- What do you want to call the rescaled image? *RescaledGFP*
- Rescaling method. (S) Stretch the image (0 to 1). (E) Enter the minimum and maximum values in the boxes below. (G) rescale so all pixels are equal to or Greater than one. (M) Match the maximum of one image to the maximum of another. (C) Convert to 8 bit. (T) Divide by loaded text value. See the help for details. *Stretch 0 to 1*
- (Method E only): Enter the intensity from the original image that should be set to the lowest value in the rescaled image, or type AA to calculate the lowest intensity automatically from all of the images to be analyzed and AE to calculate the lowest intensity from each image independently. *AA*
- (Method E only): Enter the intensity from the original image that should be set to the highest value in the rescaled image, or type AA to calculate the highest intensity automatically from all of the images to be analyzed and AE to calculate the highest intensity from each image independently. *AA*
- (Method E only): What value should pixels at the low end of the original intensity range be mapped to (range [0,1])? *0*
- (Method E only): What value should pixels at the high end of the original intensity range be mapped to (range [0,1])? *1*
- (Method E only): What value should pixels *below* the low end of the original intensity range be mapped to (range [0,1])? *0*
- (Method E only): What value should pixels *above* the high end of the original intensity range be mapped to (range [0,1])? *1*
- (Method M only): What did you call image whose maximum you want rescaled image to match? *Orig*
- (Method T only): What did you call the loaded text in the LoadText module?
-

Module #5: IdentifyPrimAutomatic

- What did you call the images you want to process? *RFP*
- What do you want to call the objects identified by this module? *Nuclei*
- Typical diameter of objects, in pixel units (Min,Max): *6,40*
- Discard objects outside the diameter range? *Yes*
- Try to merge too small objects with nearby larger objects? *No*
- Discard objects touching the border of the image? *Yes*
- Select an automatic thresholding method or enter an absolute threshold in the range [0,1]. To choose a binary image, select "Other" and type its name. Choosing 'All' will use the Otsu Global method to calculate a single threshold for the entire image group. The other methods calculate a threshold for each image individually. "Set interactively" will allow you to manually adjust the threshold during the first cycle to determine what will work well. *Otsu Global*
- Threshold correction factor *2*
- Lower and upper bounds on threshold, in the range [0,1] *0.0013,1*
- For MoG thresholding, what is the approximate fraction of image covered by objects? *0.01*
- Method to distinguish clumped objects (see help for details): *Intensity*
- Method to draw dividing lines between clumped objects (see help for details): *Intensity*
- Size of smoothing filter, in pixel units (if you are distinguishing between clumped objects). Enter 0 for low resolution images with small objects (~< 5 pixel diameter) to prevent any smoothing. *Automatic*
- Suppress local maxima within this distance, (a positive integer, in pixel units) (if you are distinguishing between clumped objects) *Automatic*
- Speed up by using lower-resolution image to find local maxima? (if you are distinguishing between clumped objects) *Yes*
- Enter the following information, separated by commas, if you would like to use the Laplacian of Gaussian method for identifying objects instead of using the above settings: Size of neighborhood (height, width), Sigma, Minimum Area, Size for Wiener Filter (height, width),Threshold *Do not use*
- What do you want to call the outlines of the identified objects (optional)? *NucleiOutline*

 Do you want to fill holes in identified objects? *Yes*
 Do you want to run in test mode where methods for distinguishing clumped objects are compared? *No*

Module #6: MeasureObjectAreaShape
 What did you call the objects that you want to measure? *Nuclei*
 Would you like to calculate the Zernike features for each object? *Yes*

Module #7: MeasureObjectIntensity
 What did you call the greyscale images you want to measure? *GFP*
 What did you call the objects that you want to measure? *Nuclei*

Module #8-15: MeasureTexture
- What did you call the greyscale images you want to measure? *GFP*
- What did you call the objects that you want to measure? *Nuclei*
- What is the scale of texture? *1-8*

Module #16: MeasureObjectIntensity
 What did you call the greyscale images you want to measure? *RFP*
 What did you call the objects that you want to measure? *Nuclei*

Module #17-24: MeasureTexture
 What did you call the greyscale images you want to measure? *RFP*
 What did you call the objects that you want to measure? *Nuclei*
 What is the scale of texture? *1-8*

Module #25: ExpandOrShrink
 What did you call the objects that you want to expand or shrink? *Nuclei*
 What do you want to call the expanded or shrunken objects? *ExpandNuclei*
 Were the objects identified using an Identify Primary or Identify Secondary module (note: shrinking results are not perfect with Secondary objects)? *Primary*
 Do you want to expand or shrink the objects? *Expand*
 Enter the number of pixels by which to expand or shrink the objects, or "Inf" to either shrink to a point or expand until almost touching, or 0 (the number zero) to simply add partial dividing lines between objects that are touching (experimental feature). *2*
 What do you want to call the outlines of the identified objects (optional)? *ExpandedNucleiOutline*

Module #26: MeasureObjectAreaShape
 What did you call the objects that you want to measure? *ExpandNuclei*
 Would you like to calculate the Zernike features for each object? *Yes*

Module #27: MeasureObjectIntensity
 What did you call the greyscale images you want to measure? *GFP*
 What did you call the objects that you want to measure? *ExpandNuclei*

Module #28-35: MeasureTexture
 What did you call the greyscale images you want to measure? *GFP*
 What did you call the objects that you want to measure? *ExpandNuclei*
 What is the scale of texture? *1-8*

Module #36: MeasureObjectIntensity
- What did you call the greyscale images you want to measure? *RFP*
- What did you call the objects that you want to measure? *ExpandNuclei*

Module #37-44: MeasureTexture

- What did you call the greyscale images you want to measure? *RFP*
- What did you call the objects that you want to measure? *ExpandNuclei*
- What is the scale of texture? *1-8*

Module #45: IdentifySecondary
 What did you call the primary objects you want to create secondary objects around? *Nuclei*
 What do you want to call the objects identified by this module? *Cells*
 Select the method to identify the secondary objects (Distance - B uses background; Distance - N does not): *Propagation*
 What did you call the images to be used to find the edges of the secondary objects? For DISTANCE - N, this will not affect object identification, only the final display. *RescaledRFP*
 Select an automatic thresholding method or enter an absolute threshold in the range [0,1]. To choose a binary image, select "Other" and type its name. Choosing 'All' will use the Otsu Global method to calculate a single threshold for the entire image group. The other methods calculate a threshold for each image individually. Set interactively will allow you to manually adjust the threshold during the first cycle to determine what will work well. *Otsu Global*
 Threshold correction factor *0.8*
 Lower and upper bounds on threshold, in the range [0,1] *0.04,1*
 For MoG thresholding, what is the approximate fraction of image covered by objects? *0.01*
 For DISTANCE, enter number of pixels by which to expand the primary objects [Positive integer] *10*
 For PROPAGATION, enter the regularization factor (0 to infinity). Larger=distance,0=intensity *0.05*
 What do you want to call the outlines of the identified objects (optional)? *CellOutline*
 Do you want to run in test mode where each method for identifying secondary objects is compared? *No*

Module #46: MeasureObjectAreaShape
 What did you call the objects that you want to measure? *Cells*
 Would you like to calculate the Zernike features for each object? *Yes*

Module #47: MeasureObjectIntensity
 What did you call the greyscale images you want to measure? *GFP*
 What did you call the objects that you want to measure? *Cells*

Module #48-55: MeasureTexture
 What did you call the greyscale images you want to measure? *GFP*
 What did you call the objects that you want to measure? *Cells*
 What is the scale of texture? *1-8*

Module #56: MeasureObjectIntensity
 What did you call the greyscale images you want to measure? *RFP*
 What did you call the objects that you want to measure? *Cells*

Module #57-64: MeasureTexture
 What did you call the greyscale images you want to measure? *RFP*
 What did you call the objects that you want to measure? *Cells*
 What is the scale of texture? *1-8*

Module #65: OverlayOutlines
 On which image would you like to display the outlines? *RescaledRFP*
 What did you call the outlines that you would like to display? *CellOutline*
 Would you like to set the intensity (brightness) of the outlines to be the same as the brightest point in the image, or the maximum possible value for this image format? *Max of image*
 What do you want to call the image with the outlines displayed? *CellRFP*
 For color images, what do you want the color of the outlines to be? *Red*

Module #66: OverlayOutlines

- On which image would you like to display the outlines? *RescaledGFP*
- What did you call the outlines that you would like to display? *ExpandedNucleiOutline*
- Would you like to set the intensity (brightness) of the outlines to be the same as the brightest point in the image, or the maximum possible value for this image format? *Max of image*
- What do you want to call the image with the outlines displayed? *ExpNucleiGFP*
- For color images, what do you want the color of the outlines to be? *Green*

Module #67: ExportToDatabase

- What type of database do you want to use? *MySQL*
- For MySQL only, what is the name of the database to use? *FociDB*
- What prefix should be used to name the tables in the database (should be unique per experiment, or leave "Do not use" to have generic Per_Image and Per_Object tables)? *Do not use*
- What prefix should be used to name the SQL files? *SQL_*
- Enter directory where the SQL files are to be saved. Type period (.) to use the default output folder. .
- Do you want to create a CellProfiler Analyst properties file? *Yes*

Module #68: CreateBatchFiles

- What is the path to the folder where the batch control file (Batch_data.mat) will be saved? Leave a period (.) to use the default output folder. .
- If pathnames are specified differently between the local and cluster machines, enter that part of the pathname from the local machine's perspective, omitting trailing slashes. Otherwise, leave a period (.) */Volumes/MetaXpress/*
- If pathnames are specified differently between the local and cluster machines, enter that part of the pathname from the cluster machines' perspective, omitting trailing slashes. Otherwise, leave a period (.) */home/MetaXpress/*

Note: This module must be the last one in the analysis pipeline.

Supplemental References

Berriz, G.F., King, O., Bryant, B., Sander, C., and Roth, F.P. (2003). Characterizing gene sets with FuncAssociate. Bioinformatics *19*, 2502-4.

Byrne, K.P., and Wolfe, K.H. (2005). The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res. *15*, 1456–61.

Chang, H.-Y., Lawless, C., Addinall, S.G., Oexle, S., Taschuk, M., Wipat, A., Wilkinson, D.J., and Lydall, D. (2011). Genome-wide analysis to identify pathways affecting telomere-initiated senescence in budding yeast. G3 (Bethesda) *1*, 197–208.

Gatbonton, T., Imbesi, M., Nelson, M., Akey, J.M., Ruderfer, D.M., Kruglyak, L., Simon, J.A., and Bedalov, A. (2006). Telomere length as a quantitative trait: genome-wide survey and genetic mapping of telomere length-control genes in yeast. PLoS Genet. *2*, e35.

Gietz, R., and Woods, R. (2002). Yeast Transformation by LiAc/SS Carrier DNA/PEG. Methods in Enzymology *350*, 1–10.

Gu, Z., Cavalcanti, A., Chen, F.-C., Bouman, P., and Li, W.-H. (2002). Extent of gene duplication in the genomes of Drosophila, nematode, and yeast. Mol. Biol. Evol. *19*, 256–262.

Haber, J.E. (2002). Uses and abuses of HO endonuclease. Methods in Enzymology *350*, 141–164.

Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. Cell *95*, 717–728.

Huang, D., Moffat, J., and Andrews, B. (2002). Dissection of a complex phenotype by functional genomics reveals roles for the yeast cyclin-dependent protein kinase Pho85 in stress adaptation and cell integrity. Mol Cell Biol. *22*, 5076-88.

Myers, C., Barrett, D., Hibbs, M., Huttenhower, C., and Troyanskaya, O. (2006). Finding function: evaluation methods for functional genomic data. BMC Genomics *7*, 187.

Ribeyre, C., and Shore, D. (2012). Anticheckpoint pathways at telomeres in yeast. Nat. Struct. Mol. Biol. *19*, 307–313.

Rost, B. (1999). Twilight zone of protein sequence alignments. Protein Eng. *12*, 85–94.

Sinclair, D.A. (1997). Accelerated Aging and Nucleolar Fragmentation in Yeast sgs1 Mutants. Science *277*, 1313–1316.

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

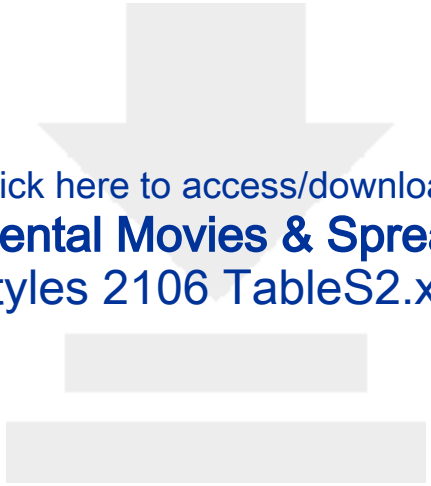
Stirling, P.C., Bloom, M.S., Solanki-Patil, T., Smith, S., Sipahimalani, P., Li, Z., Kofoed, M., Ben-Aroya, S., Myung, K., and Hieter, P. (2011). The Complete Spectrum of Yeast Chromosome Instability Genes Identifies Candidate CIN Cancer Genes and Functional Roles for ASTRA Complex Components. PLoS Genet. *7*, e1002057.

Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. Nature *449*, 54–61.

Yuen, K., Warren, C., and Chen, O. (2007). Systematic genome instability screens in yeast and their potential

relevance to cancer. Proc. Natl. Acad. Sci. USA *104*, 3925-30.

Table S2



Click here to access/download

Supplemental Movies & Spreadsheets

Styles 2106 TableS2.xls

Table S4

Click here to access/download

Supplemental Movies & Spreadsheets

Styles 2016 TableS4.xls

Table S5

Click here to access/download

Supplemental Movies & Spreadsheets

Styles 2016 TableS5.xls

Click here to access/download

Supplemental Movies & Spreadsheets

Stylesetal_MovieS1A.avi

Click here to access/download

Supplemental Movies & Spreadsheets

Stylesetal_MovieS1B.avi

Click here to access/download

Supplemental Movies & Spreadsheets

Stylesetal_MovieS1C.avi

Click here to access/download

Supplemental Movies & Spreadsheets

Stylesetal_MovieS1D.avi