Complementation of Yeast Genes with Human Genes as an Experimental Platform for Functional Testing of Human Genetic Variants

Akil Hamza,* Erik Tammpere,* Megan Kofoed,* Christelle Keong,* Jennifer Chiang,† Guri Giaever,† Corey Nislow,† and Philip Hieter*,1

*Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4, and †Department of Pharmaceutical Sciences, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3

ABSTRACT While the pace of discovery of human genetic variants in tumors, patients, and diverse populations has rapidly accelerated, deciphering their functional consequence has become rate-limiting. Using cross-species complementation, model organisms like the budding yeast, *Saccharomyces cerevisiae*, can be utilized to fill this gap and serve as a platform for testing human genetic variants. To this end, we performed two parallel screens, a one-to-one complementation screen for essential yeast genes implicated in chromosome instability and a pool-to-pool screen that queried all possible essential yeast genes for rescue of lethality by all possible human homologs. Our work identified 65 human cDNAs that can replace the null allele of essential yeast genes, including the nonorthologous pair y*RFT1*/h*SEC61A1*. We chose four human cDNAs (h*LIG1*, h*SSRP1*, h*PPP1CA*, and h*PPP1CC*) for which their yeast gene counterparts function in chromosome stability and assayed in yeast 35 tumor-specific missense mutations for growth defects and sensitivity to DNA-damaging agents. This resulted in a set of human–yeast gene complementation pairs that allow human genetic variants to be readily characterized in yeast, and a prioritized list of somatic mutations that could contribute to chromosome instability in human tumors. These data establish the utility of this cross-species experimental approach.

KEYWORDS human–yeast cross-species complementation; chromosome instability; tumor-specific missense mutations; human genetic variants; model organisms and human disease

DIFFERENTIATING passenger from driver mutations in tumor cell genomes and linking somatic variants to cancer phenotypes are major challenges. A hallmark of cancer is chromosome instability (CIN), which is an important predisposing factor in the progression and heterogeneity of tumors because it increases the likelihood of loss of tumor-suppressor genes and mutation/amplification/rearrangement of oncogenes (Lengauer *et al.* 1997; Charames and Bapat 2003; Schvartzman *et al.* 2010; Hanahan and Weinberg 2011). The budding yeast, *Saccharomyces cerevisiae*, has been used to define cellular pathways and catalog a comprehensive list of yeast genes required for the maintenance of chromosome stability (Yuen *et al.* 2007; Stirling *et al.* 2011). The yeast CIN gene list facilitates identification of candidate human CIN genes whose somatic variants may contribute to chromosome instability and tumorigenesis (Barber *et al.* 2008). Budding yeast can be exploited to screen these human genetic variants for prioritizing and directing functional studies in mammalian models (reviewed in Dunham and Fowler 2013).

There are several approaches for studying human genetic variants in yeast. Based on sequence conservation, human variants can be tested by introducing homologous mutations in the yeast ortholog. However, such sequence-based efforts restrict the number of variants that can be studied, and inferences based on this approach can be misleading as function can be maintained despite sequence divergence. For example, the impact of a nonsynonymous substitution in the context of the yeast protein may not accurately predict the same effect in the context of the human protein (Marini *et al.* 2010). A more desirable approach is functionally testing

doi: 10.1534/genetics.115.181099
Manuscript received July 26, 2015; accepted for publication September 1, 2015; published Early Online September 9, 2015.
Available freely online through the author-supported open access option.
Supporting information is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.181099/-/DC1.
[1]Corresponding author: Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, Canada V6T 1Z4. E-mail: hieter@msl.ubc.ca

variants directly in the context of the human protein sequence by utilizing cross-species complementation to "humanize" the yeast strain. This refers to the ability of a human gene (or set of genes) to complement the loss-of-function phenotype of its yeast ortholog (or set of orthologs). Cross-species complementation has previously been used in yeast to isolate human orthologs, elucidate human gene function, and characterize gene variants (reviewed in Osborn and Miller 2007; Dunham and Fowler 2013). Expression of human cDNAs in wild-type yeast cells has also been used as means to study genetic variants and characterize protein domains (Kato *et al.* 2003; Tarnowski *et al.* 2012).

While cross-species complementation can be scored by any measurable phenotype, the most straight-forward phenotype to assay and quantify is the rescue of growth defects. Complementation of essential yeast genes by candidate human gene orthologs can be tested by the ability of a human cDNA to rescue lethality caused by (i) a null allele (deletion in a haploid strain), (ii) a conditional allele under restrictive conditions (*e.g.*, temperature-sensitive strain), or (iii) down-regulation by a repressible promoter (*e.g.*, Tet system) (Kachroo *et al.* 2015). Nonessential yeast gene mutants can be assayed when they present a phenotype (*e.g.*, drug sensitivity) or the nonessential yeast gene can be converted to an essential gene by disrupting a synthetic lethal partner (Greene *et al.* 1999). Hundreds of studies testing individual genes have revealed >200 successful human–yeast complementation pairs (Dunham and Fowler 2013). Two studies systematically tested for human–yeast complementation pairs. One screened for rescue of lethality caused by inducible loss-of-function of 25 essential yeast genes (repressible promoter) following transformation of a human cDNA library (*i.e.*, pool-to-one screens) and identified six essential genes that were rescued by a human ortholog (Zhang *et al.* 2003). More recently, 176/414 essential yeast genes (as null and/or Ts mutations and/or using a repressible promoter) were found to be replaceable by their 1:1 human ortholog (*i.e.*, one-to-one screens) (Kachroo *et al.* 2015).

We report the screening of 621 essential yeast gene null mutants for complementation by all potential human homologs in two parallel screens ("one-to-one" screens for 199 essential CIN genes corresponding to 322 candidate complementation pairs, and a "pool-to-pool" screen for all possible essential yeast genes). In addition, we demonstrate the feasibility of using yeast to test human gene variants by extending cross-species complementation to 35 tumor-specific mutations in genes functioning in chromosome stability.

Materials and Methods

Generating expression vectors for the complementation screens

One-to-one screen: Human cDNAs in Gateway-compatible entry clones were obtained from hORFeome V8.1 (Yang *et al.* 2011) and shuttled into yeast destination vectors (Alberti *et al.* 2007) using LR clonase II (Invitrogen) to generate expression clones. The destination vector used was pAG416GPD-ccdB-HA (*URA3*, CEN, constitutive GPD promoter, C-terminal HA tag) with a stop codon contributed by the vector backbone resulting in a 55-amino-acid C-terminal extension. The identity of the human cDNA was confirmed by sequencing the expression vector using a common primer that hybridizes to the vector backbone (CAGGAAACAGCTATGAC).

Pool-to-pool screen: The same experimental outline was followed as the one-to-one screen with the following modifications and additional steps. Human cDNAs were randomly grouped into 13 pools with each pool comprising up to 96 unique entry clones. Each of the 13 pools of entry clones were shuttled en masse into a yeast destination vector to generate 13 pools of expression vectors (Arnoldo *et al.* 2014). The destination vector used was pAG416GPD-ccdB (*URA3*, CEN, constitutive GPD promoter) with a stop codon contributed by the vector backbone resulting in a 50-amino-acid C-terminal extension. To ensure sufficient coverage of each expression vector within a pool, the bulk LR reaction was repeated three times to obtain a minimum of 10,000 transformants for 100-fold coverage (~96 entry clones × 100).

Yeast strains and media used in complementation screens

One-to-one screen: Generated expression vectors were transformed into the corresponding haploid-convertible heterozygous diploid knockout yeast strain (Pan *et al.* 2004) and transformants were selected on −Ura media. Transformants were then inoculated in liquid sporulation media (1% w/v potassium acetate, 0.005% w/v zinc acetate, and 0.3 mM histidine) (Pan *et al.* 2007) to a cell density of ~1–2 OD_{600} at 25° with shaking for 5 days and sporulation efficiency was assessed using microscopy. Following sporulation, 50 μl of cells were resuspended in 1 ml water of which 100 μl was plated on the haploid selection media MM −Ura (−Leu −His −Arg −Ura +canavanine +G418) (Pan *et al.* 2007) and incubated at 30°. To confirm that the generated yeast essential haploid knockout is dependent on the expression vector, cells were then replica plated on MM +5-FOA (0.1%) (−Leu −His −Arg +canavanine +G418 +5-FOA) and incubated at 30°.

Pool-to-pool screen: The same experimental outline was followed as the one-to-one screen except with the following modifications. Heterozygous yeast strains were pinned in a 96-well array format on YPD +G418 (200 μg/ml) agar plates and incubated at 30°. Colonies were then scraped and pooled (by suspension in 1 ml YPD), and inoculated in 250 ml YPD and allowed to grow for only two generations to prevent competitive outgrowth before proceeding with the transformation (Pan *et al.* 2007) for protocol on high-efficiency transformation). The 250 ml yeast culture was then divided into 13 equal aliquots into which 13 pools of expression vector DNAs were transformed. Creating pools of transforming DNA and cells to be transformed was done to ensure

equal representation of yeast strains across all pools. To ensure sufficient coverage of each yeast strain/vector combination in sufficient numbers, the transformation was repeated a second time for each pool to obtain a minimum of 6 million transformants for 100-fold coverage (621 yeast strains × ~96 expression vectors × 100). Transformed diploid colonies were then scraped, pooled, and inoculated in 13 separate 50 ml sporulation cultures as described. Sporulated cultures were then pelleted, resuspended in water, and plated on MM −Ura for haploid selection, after which the haploid-converted cells were scraped, pooled, and re-plated again on MM −Ura to obtain single colonies, which were confirmed by replica plating to MM +5-FOA. For each pool, ~20–50 5-FOA-sensitive colonies were isolated. To determine the identity of the yeast strain, genomic DNA was prepped and the yeast barcode was amplified using the U1/D1 primers to allow sequencing of the UPTAG/DNTAG using the kanB or kanC primers (Giaever *et al.* 2002). To determine the identity of the rescuing human cDNA, expression vectors were isolated and sequenced using a common primer that hybridizes to the vector backbone (CAGGAAACAGCTATGAC). Each potential hit (a rescued yeast colony) was reconfirmed by retransformation of the extracted plasmid into the corresponding heterozygous diploid as described for the one-to-one screen. To confirm the nonorthologous hits, both extracted pAG416GPD-h*SEC61A1* and newly generated pAG416GPD-h*SEC61A1*-HA expression vectors along with pAG416GPD-h*RFT1*-HA and GAL-inducible y*SEC61* from the FLEX array (Hu *et al.* 2007) were separately transformed into the *RFT1/rft1*Δ heterozygous diploid yeast strain and the experiment was carried out as described for the one-to-one screen.

Generating lists and analysis for the complementation screens

A comprehensive list of essential yeast CIN genes was obtained from Stirling *et al.* (2011), while a list of essential yeast genes was obtained from the Yeast Deletion Project (Giaever *et al.* 2002). A list of human genes was generated from two sources: Yeastmine (Balakrishnan *et al.* 2012) and Ensembl BioMart (Kinsella *et al.* 2011) databases. Gene Ontology (GO) processes and other features used to analyze the complementation set were obtained from Yeastmine and each feature was represented as a proportion of the total number of genes input for each gene set. Significance for each feature was calculated using the hypergeometric distribution and subjected to the Bonferroni correction to obtain the adjusted *P*-value (critical value: 0.05; number of tests: 24 for features analyzed in Figure 3, A–F). Sequence identity (%) in relation to the yeast gene was calculated for all possible human–yeast gene pairs using the Ensembl BioMart database. Significance between the % sequence identity of complementation pairs and % sequence identity of all human–yeast pairs included in this study was calculated using a Mann–Whitney test. For some manually curated orthologs and others generated from Yeastmine, sequence identity was determined using NWalign (Y. Zhang, http://zhanglab.ccmb.med.umich.edu/NW-align).

Generating variants and yeast strains for growth curves

Expression vectors and yeast strain construction: As described, h*LIG1*, h*SSRP1*, h*PPP1CA*, and h*PPP1CC* in entry clones were shuttled into the yeast destination vectors pAG416GPD-ccdB-HA (*URA3*) and pAG415GPD-ccdB-HA (*LEU2*) to generate expression clones (Supporting Information, Table S1). Missense mutations were introduced in vector pAG415GPD-hORF-HA (*LEU2*) using the QuikChange Site-Directed Mutagenesis Kit (Agilent), and verified by Sanger sequencing. To generate the yeast haploid knockout strains, tetrads were dissected following transformation of the expression vector pAG416GPD-hORF-HA (*URA3*) to the corresponding heterozygous diploid deletion strain and then haploids were maintained on −Ura media (Table S1). To generate the strains used for growth curve analysis, expression clones containing wild-type and mutant human cDNA on pAG415GPD-hORF-HA (*LEU2*) were first transformed into both the generated haploid knockouts (covered by wild-type human cDNA marked by *URA3*) and wild-type strain (BY4742) (Brachmann *et al.* 1998) and maintained on −Ura −Leu or −Leu media, respectively. For the plasmid shuffle, strains were then plated on −Leu +Ura +5-FOA media and individual colonies were picked and thereafter maintained on −Leu media. To confirm that the *URA3*-marked plasmid was lost, strains were streaked on −Ura media to observe lack of growth.

Growth curves: Unless otherwise indicated, all growth conditions were carried out in −Leu media at 30°. Strains were inoculated overnight, diluted to a cell density of 0.1 OD_{600}, and then grown to mid-log phase. Strains were then diluted to OD_{600} = 0.1 in 200 μl −Leu media with no drug, 0.01% MMS, or 100 mM HU in 96-well plates. OD_{600} readings were measured every 30 min over a period of 48 hr in a TECAN M200 plate reader. Prior to each reading, plates were shaken for 10 min. Each strain was tested in three to four replicates per plate per condition and area under the curve (AUC) value was calculated for each replicate independently. For growth curves in −Leu and −Leu +100 mM HU, AUC values were calculated from 0–24 hr corresponding to when most strains reached saturation. For growth curves in −Leu +0.01% MMS, AUC values were calculated from 0 to 48 hr. For each mutant, strain fitness was defined as the AUC of the mutant curves relative to the AUC of the wild-type allele grown on the same plate in the same media condition. Significant differences in growth in the "no drug" condition was determined against the wild-type allele in the same condition, while significant differences in growth in the "drug" condition was determined against the same allele in the no drug condition using a Student's *t*-test. Sets of all individual growth curves are shown in Figure S1, Figure S2, Figure S3, and Figure S4.

Data availability

Plasmids and strains are available upon request. Plasmids and strains used for screening tumor-specific variants in yeast are listed in Table S1.

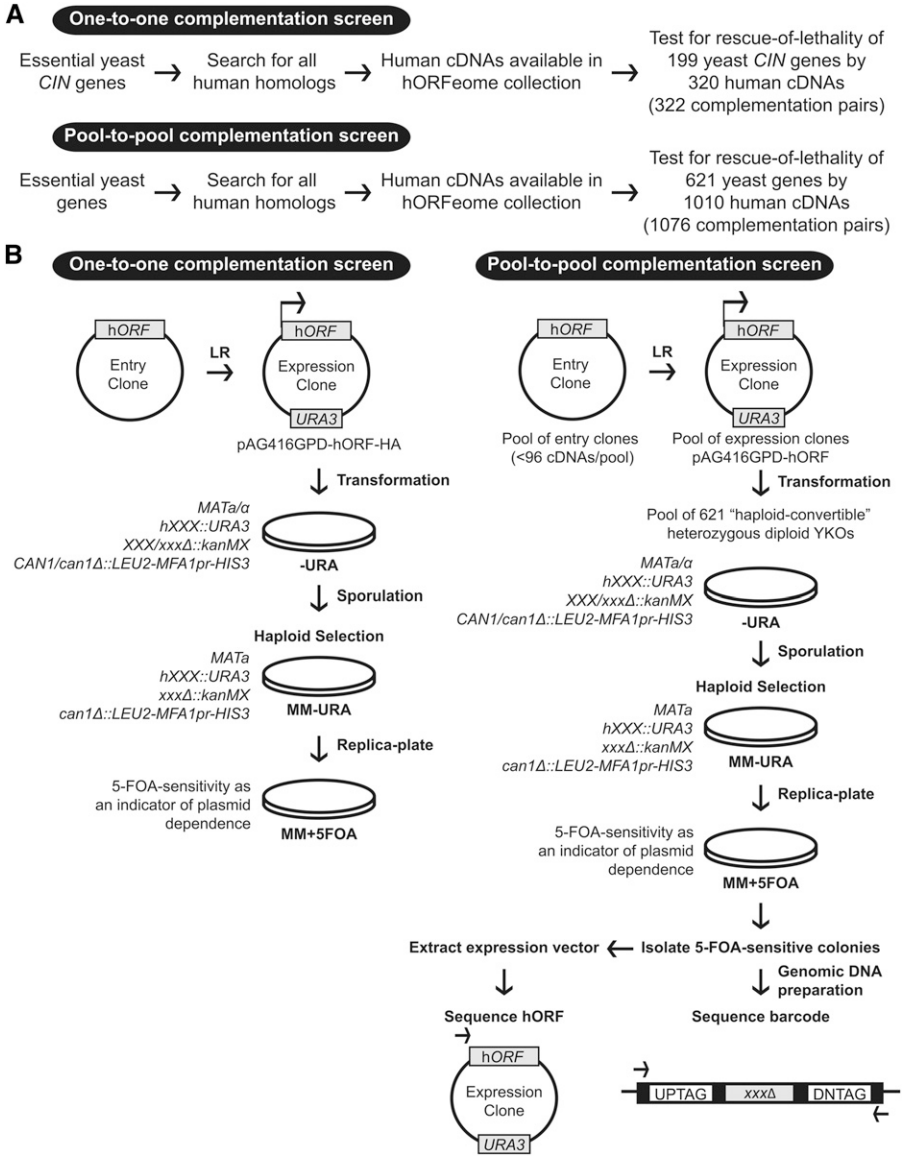

Figure 1 Overview of the experimental design for the complementation screens. (A) Pipeline outlining which human–yeast complementation pairs were included in both screens. (B) Flowchart for the complementation screens. Human cDNAs were shuttled from entry clones to indicated yeast destination vectors to generate yeast expression vectors. Single or pooled expression vectors were then transformed to matched or pooled haploid convertible heterozygous diploids and maintained on −Ura media. Following sporulation, heterozygous diploids were plated on haploid selection media (MM −Ura). "Rescued" haploids were tested for plasmid dependency by replica plating on MM +5-FOA. For the pooled screen, 5-FOA-sensitive colonies were isolated for sequencing of yeast barcode and expression vectors.

Results

Systematic identification of human–yeast cross-species complementation pairs

A comprehensive list of yeast CIN genes revealed that many essential genes are mutable to a CIN phenotype (323 CIN genes/∼1100 essential genes, or ∼29%) (Stirling *et al.* 2011). To establish an experimental platform for testing tumor-specific somatic mutations in candidate human CIN genes, we tested all possible yeast CIN gene null mutants for complementation by all potential human homologs in a series of one-to-one screens. While one-to-one screening reduces the number of false negatives that can arise as an artifact of the skewed representation associated with pooled screening, it does not allow identification of unexpected or nonorthologous complementation pairs. Therefore, we set up a pool-to-pool screen to test all possible essential (CIN and non-CIN) yeast genes for complementation by all sequence-related human homologs that are available as full-length cDNA clones. For the purposes of this study, we use "homolog" as an umbrella term that encompasses orthologs (genes derived from speciation and that typically perform equivalent functions) and paralogs (genes related by duplication and that generally perform biologically distinct yet mechanistically related functions) (reviewed in Koonin 2005).

To set up both screens (Figure 1), we used Gateway cloning to enable the systematic shuttling of DNA fragments between cloning vectors (Hartley *et al.* 2000). Briefly, a human cDNA of interest flanked by recombination sites on an entry clone can be transferred to Gateway-compatible destination vectors in a single-step reaction to create yeast expression vectors. We used the human ORFeome V8.1, which is a Gateway-compatible clone library of sequence-confirmed human cDNAs (Yang *et al.* 2011) as the source of human ORFs. We also used the Gateway-compatible library of yeast expression vectors, which offers a selection of promoters (constitutive

Table 1 Human cDNAs that complement essential yeast deletion mutants

Yeast systematic name	Yeast standard name	Human Entrez Gene ID	Human standard name	Yeast gene brief description[d]
YBL020W	*RFT1*[a]	29927	*SEC61A1*	Membrane protein required for translocation of Man5GlcNac2-PP-Dol
YBR002C	*RER2*	79947	*DHDDS*	*Cis*-prenyltransferase involved in dolichol synthesis
YBR029C	*CDS1*	8760	*CDS2*	Phosphatidate cytidylyltransferase (CDP-diglyceride synthetase)
YBR109C	*CMD1*[b]	801; 805; 808	*CALM1*; *CALM2*; *CALM3*	Calmodulin
YBR160W	*CDC28*[b]	983; 1017	*CDK1*; *CDK2*	Cyclin-dependent kinase (CDK) catalytic subunit
YBR252W	*DUT1*	1854	*DUT*	Deoxyuridine triphosphate diphosphatase (dUTPase)
YCR012W	*PGK1*[b]	5230; 5232	*PGK1*[c]; *PGK2*	3-Phosphoglycerate kinase
YDL045C	*FAD1*[b]	80308	*FLAD1*[c]	Flavin adenine dinucleotide (FAD) synthetase
YDL064W	*UBC9*[b]	7329	*UBE2I*	SUMO-conjugating enzyme involved in the Smt3p conjugation pathway
YDL120W	*YFH1*	2395	*FXN*	Mitochondrial matrix iron chaperone
YDL147W	*RPN5*[b]	5718	*PSMD12*[c]	Subunit of the CSN and 26S proteasome lid complexes
YDL164C	*CDC9*[b]	3978	*LIG1*[c]	DNA ligase found in the nucleus and mitochondria
YDL205C	*HEM3*	3145	*HMBS*	Porphobilinogen deaminase
YDR050C	*TPI1*[b]	7167	*TPI1*	Triose phosphate isomerase, abundant glycolytic enzyme
YDR086C	*SSS1*	23480	*SEC61G*	Subunit of the Sec61p translocation complex (Sec61p-Sss1p-Sbh1p)
YDR208W	*MSS4*[b]	8394; 8395	*PIP5K1A*; *PIP5K1B*[c]	Phosphatidylinositol-4-phosphate 5-kinase
YDR236C	*FMN1*	55312	*RFK*	Riboflavin kinase, produces riboflavin monophosphate (FMN)
YDR404C	*RPB7*[b]	5436	*POLR2G*	RNA polymerase II subunit B16
YDR454C	*GUK1*	2987	*GUK1*	Guanylate kinase
YDR510W	*SMT3*[b]	7341	*SUMO1*[c]	Ubiquitin-like protein of the SUMO family
YEL026W	*SNU13*[b]	4809	*NHP2L1*	RNA binding protein
YEL058W	*PCM1*	5238	*PGM3*	Essential N-acetylglucosamine-phosphate mutase
YER094C	*PUP3*[b]	5691	*PSMB3*	Beta 3 subunit of the 20S proteasome
YER112W	*LSM4*	25804	*LSM4*	Lsm (Like Sm) protein
YER133W	*GLC7*[b]	5499; 5501	*PPP1CA*; *PPP1CC*	Type 1 serine/threonine protein phosphatase catalytic subunit
YER136W	*GDI1*	2665	*GDI2*	GDP dissociation inhibitor
YFL017C	*GNA1*[b]	64841	*GNPNAT1*[c]	Glucosamine-6-phosphate acetyltransferase
YGL001C	*ERG26*	50814	*NSDHL*	C-3 sterol dehydrogenase
YGL030W	*RPL30*[b]	6156	*RPL30*	Ribosomal 60S subunit protein L30
YGL048C	*RPT6*	5705	*PSMC5*	ATPase of the 19S regulatory particle of the 26S proteasome
YGR024C	*THG1*	54974	*THG1L*	tRNAHis guanylyltransferase
YGR075C	*PRP38*	55119	*PRPF38B*	Unique component of the U4/U6.U5 tri-snRNP particle
YGR175C	*ERG1*	6713	*SQLE*	Squalene epoxidase
YGR185C	*TYS1*	8565	*YARS*	Cytoplasmic tyrosyl-tRNA synthetase
YGR277C	*CAB4*	80347	*COASY*	Subunit of the CoA-synthesizing protein complex (CoA-SPC)
YGR280C	*PXR1*	54984	*PINX1*	Essential protein involved in rRNA and snoRNA maturation
YIL083C	*CAB2*	79717	*PPCS*	Subunit of the CoA-synthesizing protein complex (CoA-SPC)
YJL097W	*PHS1*	201562	*PTPLB*	Essential 3-hydroxyacyl-CoA dehydratase of the ER membrane
YJR006W	*POL31*[b]	5425	*POLD2*	Subunit of DNA polymerase delta (polymerase III)
YKL013C	*ARC19*[b]	10093	*ARPC4*[c]	Subunit of the ARP2/3 complex
YKL024C	*URA6*[b]	51727	*CMPK1*[c]	Uridylate kinase
YKL033W	*TTI1*[b]	9675	*TTI1*	Subunit of the ASTRA complex, involved in chromatin remodeling
YKL035W	*UGP1*[b]	7360	*UGP2*[c]	UDP-glucose pyrophosphorylase (UGPase)
YKL145W	*RPT1*	5701	*PSMC2*	ATPase of the 19S regulatory particle of the 26S proteasome
YKL189W	*HYM1*	51719; 81617	*CAB39*; *CAB39L*	Component of the RAM signaling network
YML069W	*POB3*[b]	6749	*SSRP1*	Subunit of the heterodimeric FACT complex (Spt16p-Pob3p)
YML077W	*BET5*[b]	58485	*TRAPPC1*	Core component of transport protein particle (TRAPP) complexes I-III
YMR208W	*ERG12*	4598	*MVK*	Mevalonate kinase
YMR308C	*PSE1*[b]	3843	*IPO5*	Karyopherin/importin that interacts with the nuclear pore complex
YMR314W	*PRE5*[b]	5682	*PSMA1*	Alpha 6 subunit of the 20S proteasome
YOL133W	*HRT1*	9978	*RBX1*	RING-H2 domain core subunit of multiple ubiquitin ligase complexes
YOR143C	*THI80*	27010	*TPK1*	Thiamine pyrophosphokinase
YOR149C	*SMP3*[b]	80235	*PIGZ*	Alpha 1,2-mannosyltransferase
YOR176W	*HEM15*	2235	*FECH*	Ferrochelatase

(continued)

Table 1, *continued*

Yeast systematic name	Yeast standard name	Human Entrez Gene ID	Human standard name	Yeast gene brief description[d]
YOR236W	*DFR1*	1719	*DHFR*	Dihydrofolate reductase involved in tetrahydrofolate biosynthesis
YPL117C	*IDI1*[b]	3422	*IDI1*[c]	Isopentenyl diphosphate:dimethylallyl diphosphate isomerase
YPR082C	*DIB1*[b]	10907	*TXNL4A*[c]	17-kDa component of the U4/U6aU5 tri-snRNP
YPR113W	*PIS1*	10423	*CDIPT*	Phosphatidylinositol synthase

[a] Complementation by nonorthologous gene.
[b] Yeast CIN gene.
[c] Complementation identified by both screens.
[d] Brief description obtained from Yeastmine.

GPD *vs.* inducible GAL), plasmid copy number (2μ *vs.* CEN backbone), N- or C-terminal tags, and auxotrophic markers (Alberti *et al.* 2007). To express human cDNAs in yeast, we selected a centromere-based backbone and a constitutive promoter to reduce plasmid copy-number variation and minimize potential toxicity that can arise from overexpression of human cDNAs in yeast (Tugendreich *et al.* 2001). To test for complementation of essential gene null mutations, we assayed for rescue of lethality of the haploid yeast gene deletion following sporulation of the "haploid-convertible" heterozygous diploids (Pan *et al.* 2004).

To design and execute the complementation assay, we compiled all sequence homologs of essential yeast genes generated from multiple sources using the Yeastmine database (Balakrishnan *et al.* 2012), Ensembl BioMart database (Kinsella *et al.* 2011) and to a lesser extent through manual curation. Each candidate complementation pair (human cDNA and cognate yeast gene) was included in our screens based on the availability and quality control of both the human cDNA in the hORFeome collection and the corresponding yeast strain in the heterozygous collection. In an effort to be as comprehensive as possible, our screens tested all possible complementation pairs even when the least diverged candidate ortholog was not available in the hORFeome collection. For instance, y*CDC42* is a rho-like GTPase required for the establishment of cell polarity (Johnson 1999). The list of homologs for y*CDC42* included genes such as h*CDC42*, h*RHOJ*, h*RHOV*, h*RHOU*, and other related members of the RHO family of small GTPases. Of these homologs, the least diverged ortholog, h*CDC42* (80% sequence identity, previously shown to complement the yeast deletion mutant (Kachroo *et al.* 2015), was not available in the human ORFeome V8.1 collection and therefore not tested in this study. We did test h*RHOJ* (60% sequence identity) and determined that it does not complement the null allele of y*CDC42*. In total, we tested 199 essential CIN deletion mutants for rescue of lethality by 320 human cDNAs one-by-one (322 candidate complementation pairs) and 621 essential gene deletion mutants for rescue of lethality by 1010 human cDNAs pool-to-pool (1076 candidate complementation pairs) (Figure 1A, Table S2 and Table S3). Our screens identified 65 human cDNAs that complement the null allele of 58 essential yeast genes, including a complementation pair of nonorthologous proteins (yeast Rft1 and human Sec61A1) (Table 1). When compared to a curated list of complementation pairs available from the *Saccharomyces* Genome Database (SGD) that includes the recent results of Kachroo *et al.* (2015) (http://downloads.yeastgenome.org/curation/literature/functional_complementation.tab (June 2015) and other published reports, our work identified 20 novel complementation pairs in which the human cDNA rescues lethality of the yeast null allele (Table S4).

Assessing features of essential yeast genes that predict replaceability

We used Yeastmine to assess GO terms and looked for features that predict replaceability of essential yeast genes. Our yeast complementation set ($n = 58$) is composed predominantly of metabolic genes (Figure 2), a feature also observed in a recent large-scale complementation study (Kachroo *et al.* 2015). In general, the complemented yeast genes encode proteins that localize to the cytoplasm and associated organelles rather than the nucleus or other nuclear-associated regions (Figure 3A). When grouped by molecular function, the complementation set is enriched for proteins that display catalytic activity and other enzymatic features such as nucleotide, metal ion, and cofactor binding (Figure 3B).

Even though the complemented yeast genes display no difference in the number of genetic interactions (Figure 3C), there is a marked difference in the number of physical interactions in this set: replaceable yeast proteins tend to have fewer physical interactions (Figure 3D) and are less likely to be part of macromolecular complexes (Figure 3E). We further show that replaceable yeast genes tend to be short (Figure 3F), corroborating the observation of Kachroo *et al.* (2015).

Complementation pairs tend to have higher than average sequence identity/similarity, but sequence identity alone is a poor predictor of replaceability as evidenced by the observations that high sequence similarity does not guarantee replaceability and complementation pairs can have low sequence similarity (Figure 3G). We assessed whether replaceability by one human homolog predicts the same outcome for additional human homologs. In one case, we found that the glycolytic enzyme y*PGK1* can be replaced by either h*PGK1* (66% sequence identity) or h*PGK2* (63% sequence identity),

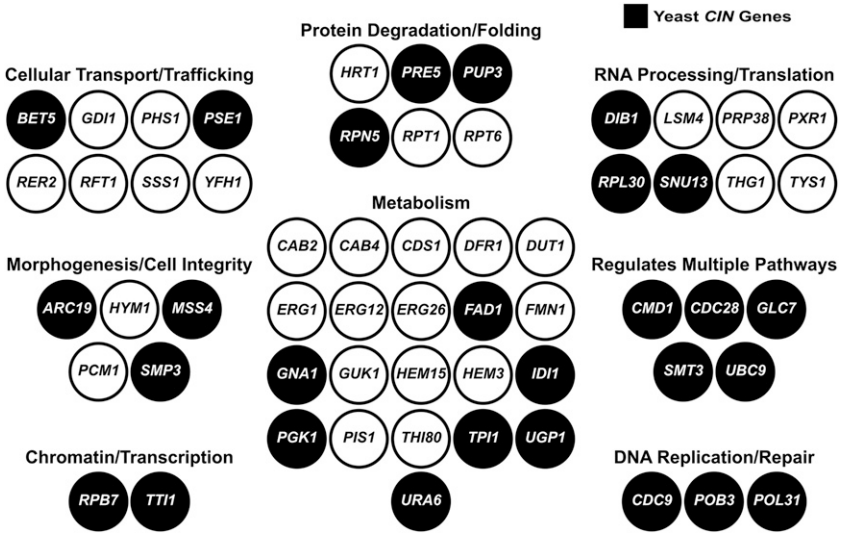

Figure 2 Essential yeast genes that are replaceable by human genes. A total of 58 yeast genes are represented by nodes and grouped according to cellular processes (Yeastmine). Yeast CIN genes are represented by black nodes.

human proteins that share 87% sequence identity and catalyze the same reaction but are differentially expressed (McCarrey and Thomas 1987). In the only other example, the phosphatase y*GLC7* can be replaced by the isozymes h*PPP1CA* (84% sequence identity) and h*PPP1CC* (84% sequence identity), human proteins that are 91% identical and part of the highly conserved PP1 subfamily of protein phosphatases (Ceulemans and Bollen 2004). In contrast, some yeast genes such as y*DIB1*, y*IDI1*, and y*SMT3* are only replaceable by one of multiple human homologs, whereas others like y*CMD1*, y*MSS4*, and y*CDC28* are replaceable by several (but not all) human homologs. For these cases and the previously mentioned y*CDC42*, we observed that replaceable human proteins share higher sequence identity with the yeast protein than the nonreplaceable ones (Table S5), suggesting that the least diverged homologs were the most likely to complement. The only contradiction of this observation is h*SEC61A1*, which can replace the nonorthologous y*RFT1* and rescue lethality of *rft1*Δ. Sec61 forms an ER membrane channel and is required for cotranslational and post-translational translocation of proteins into the ER (Osborne *et al.* 2005), while Rft1 also functions at the ER membrane and is implicated in the translocation of lipid-linked oligosaccharides into the ER (Helenius *et al.* 2002). We further demonstrated that h*RFT1* cannot complement y*RFT1* (also shown by Kachroo *et al.* 2015), and that overexpression of the yeast ortholog of h*SEC61A1*, y*SEC61*, also fails to rescue lethality of *rft1*Δ (Figure 4). The fact that h*SEC61A1* also fails to complement y*SEC61* (Table S2) highlights the unexpected and complex manner in which the human protein acts to functionally substitute for the nonorthologous yeast protein. Overall, our observations and those of Kachroo *et al.* (2015) suggest that complementation cannot be predicted by sequence similarity.

Characterizing tumor-specific variants using cross-species complementation

Human–yeast complementation enables direct testing of human gene variants for function. Genetic variants can be screened rapidly and are characterized in the context of the human protein. For example, complementation of yeast *cys4*Δ by its human ortholog *CBS* gene enabled the testing of 84 variants from patients with homocystinuria for functionality and cofactor dependence (Mayfield *et al.* 2012). One resource from our human–yeast complementation screen is a list of human genes whose variants can be characterized in a yeast deletion strain. We focused on testing tumor-specific mutations found in human orthologs of yeast CIN genes. Our complementation set (Table 1) includes 28 essential yeast CIN genes rescued by 34 candidate human CIN genes. We observe that the replaceable yeast CIN genes function across diverse biological processes, including those pathways predicted to protect genome integrity (*e.g.*, DNA replication/repair and chromatin-related processes) and more peripheral pathways (*e.g.*, metabolism and cellular trafficking) (Figure 2).

We selected three yeast CIN genes corresponding to four complementation pairs (y*CDC9*:h*LIG1*, y*POB3*:h*SSRP1*, y*GLC7*:h*PPP1CA*, and y*GLC7*:h*PPP1CC*) and assessed the impact of single amino acid substitutions in the human protein. Briefly, y*CDC9* encodes DNA ligase, an essential enzyme that joins Okazaki fragments during DNA replication and also functions in DNA repair pathways (Lindahl and Barnes 1992); y*POB3* is a subunit of the heterodimeric FACT complex (y*POB3*-y*SPT16*) that functions in nucleosome reorganization to facilitate transcription and replication processes (Formosa 2008); y*GLC7* is the catalytic subunit of protein phosphatase which, depending on its bound regulatory subunit, regulates numerous cellular processes (Cannon 2010). We selected one cancer type (colorectal cancer) and screened all reported missense mutations compiled in two databases: Catalogue of Somatic Mutations in Cancer (COSMIC) (Forbes *et al.* 2015) and cBIOPORTAL for Cancer Genomics (Cerami *et al.* 2012; Gao *et al.* 2013). In total, 35 single amino acid substitutions were constructed corresponding to 16 in h*LIG1*, 12 in h*SSRP1*, 2 in h*PPP1CA*, and 5 in h*PPP1CC*. As a control, we expressed each human gene variant in a wild-type yeast

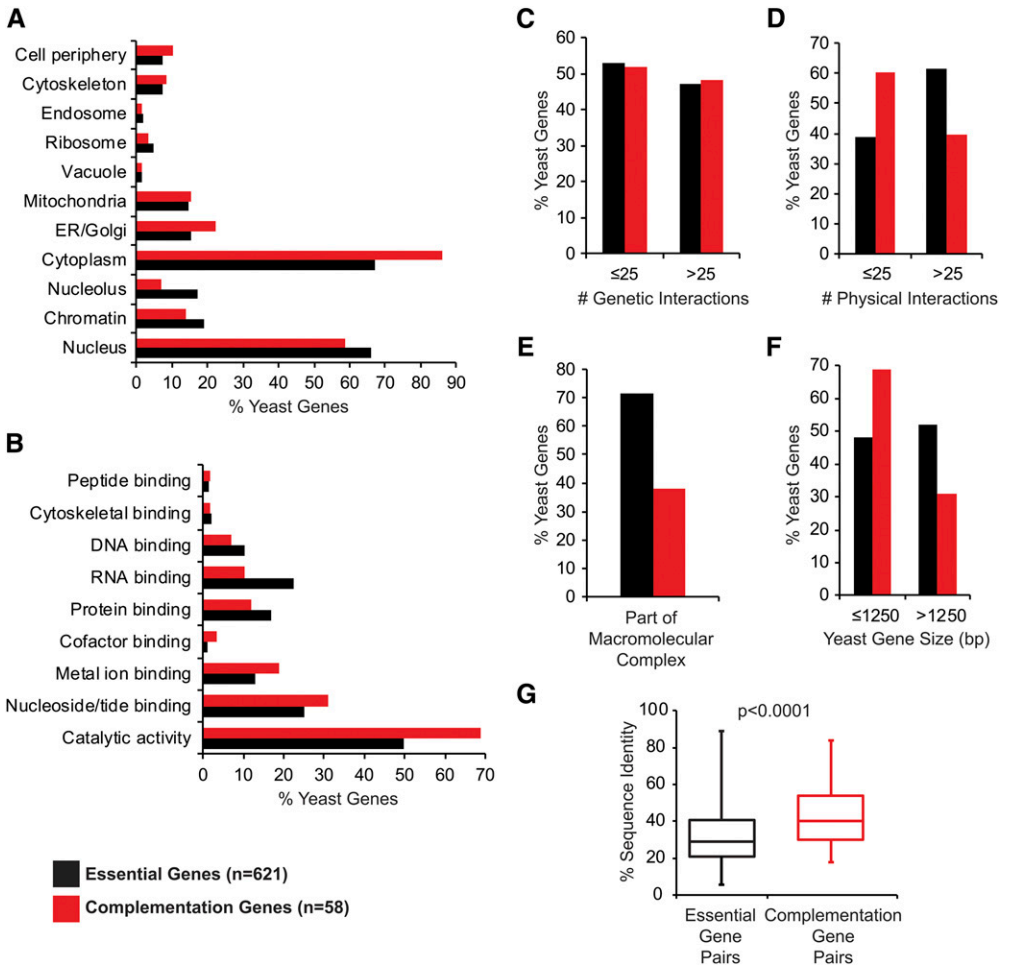

Figure 3 Analyzing features of essential yeast genes that predict replaceability including (A) localization patterns, (B) molecular function, (C) no. of genetic interactions, (D) no. of physical interactions, (E) part of macromolecular complexes, (F) yeast gene size, and (G) human–yeast sequence identity. Localization data, Gene Ontology (GO) terms, no. of genetic/physical interactions, and gene size for each yeast gene were obtained from Yeastmine and each feature is represented as a proportion of the total number of genes input for each set (n = 621 for all essential genes included in both screens and n = 58 for the complementation genes). Overall, the complementation set was enriched for yeast proteins that localize to the cytoplasm (P = 8.18E-03), have catalytic activity (P = 2.28E-02), less physical interactions (P = 5.77E-03), are less likely to be part of macromolecular complexes (P = 3.36E-07), and have smaller gene size (P = 9.16E-03). For sequence identity, "essential gene pairs" refers to the 1076 human–yeast pairs included in this study corresponding to 621 yeast genes and "complementation gene pairs" refers to the 65 complementation pairs corresponding to 58 yeast genes. The box plot highlights the median and range of sequence identity for each set of gene pairs.

strain and compared the growth of these yeast strains to the growth of the same yeast strain expressing the wild-type human allele. This step is required to identify any mutants whose ectopic expression may impact the growth phenotype of the yeast strain either by causing toxicity to the yeast cell or increasing the growth rate. Of the 35 mutants analyzed, the K228E mutation in h*SSRP1* is the only amino acid substitution whose ectopic expression decreased fitness of wild-type yeast (Table S6, Figure S1).

We used a plasmid shuffle strategy to introduce the human gene variants in the corresponding yeast gene deletion mutants. Haploid yeast gene knockout strains carrying the rescuing human cDNA on a *URA3*-marked plasmid were transformed with wild-type or mutated human cDNA on a *LEU2*-marked plasmid and plated on media containing 5-FOA (Figure 5), which selects for cells that have lost the *URA3* plasmid. Growth on 5-FOA plates indicates that the mutant human cDNA is able to complement the essential gene. Lack of growth on 5-FOA indicates that the mutation in the human cDNA results in a nonfunctional protein. We observed that 4 of the 35 mutant alleles [2 in h*SSRP1* (K228E, S481P), 1 in h*PPP1CA* (Y272C), and 1 in h*PPP1CC* (L289R)] are nonfunctional (Table 2). Although h*SSRP1*-K228E was found to be detrimental in the context of the human gene, the same substitution in the conserved site of the yeast protein (yPob3-K271E) is tolerated (VanDemark *et al.* 2006). This corroborates previous observations that amino acid substitutions in the context of the yeast gene (or other orthologs) do not necessarily have the same effect in the context of the human gene (Marini *et al.* 2010), emphasizing the importance of characterizing human variants in their native gene context. To determine the accuracy of our experimental readout, we surveyed the literature and found the conserved Y272 and L289 in PP1 phosphatases are essential for substrate binding (Egloff *et al.* 1995, 1997; Zhang *et al.* 1996).

The remaining 31 mutant alleles were assessed for their ability to rescue growth of the null mutant (Table 2, Figure S2, Figure S3, and Figure S4). We observed that 19 of the 31 mutations cause a significant decrease in strain fitness relative to the wild-type allele, while 1 mutant allele (h*PPP1CC*-L201F) grows considerably better than the corresponding wild-type allele. Notably, 16 of the 19 "slow growers" are substitutions in h*LIG1*, with the mutations dispersed throughout the protein. Given that mutants with defects in chromosome stability are often sensitive to DNA-damaging agents, we assayed our 31 mutants for growth in the presence of methyl methane sulfonate (MMS) and hydroxyurea (HU).

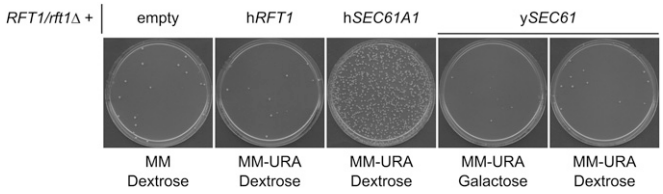

Figure 4 h*SEC61A1* complements y*RFT1*. Expression vectors (h*RFT1* and h*SEC61A1* under the control of the GPD constitutive promoter and y*SEC61* under the control of the GAL-inducible promoter) were transformed into the *RFT1/rft1*Δ heterozygous diploid yeast strain and plated on haploid selection media (MM −Ura) following sporulation. Complementation was scored by higher than background growth on MM −Ura as shown and confirmed by testing for plasmid dependency on MM +5-FOA and tetrad dissection.

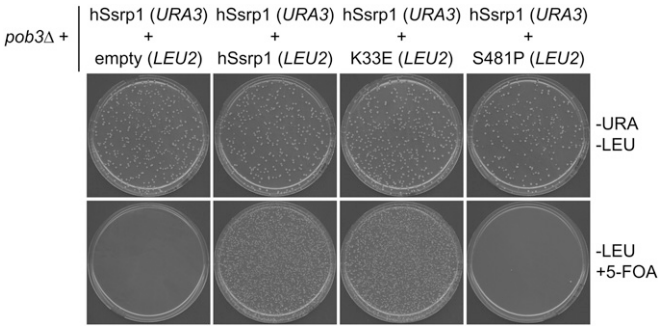

Figure 5 Plasmid shuffle used to generate strains expressing human cDNAs with missense mutations. Yeast haploid knockout strains covered by wild-type human cDNAs on *URA3*-marked vectors were transformed with the following *LEU2*-marked vectors (empty, wild-type human cDNA, and human cDNA with missense mutation) and maintained on −Ura −Leu media. Strains were plated on −Leu +5-FOA media to generate haploid yeast knockouts covered by *LEU2*-marked vectors. Strains were confirmed to have lost the *URA3*-marked plasmid by streaking on −Ura media to observe no growth. In the presented example, hSsrp1-K33E is able to complement *pob3*Δ, but S481P results in a nonfunctional hSsrp1 protein.

MMS is an alkylating agent (Beranek 1990) while HU slows DNA replication by decreasing the rate of dNTP synthesis (Koc *et al.* 2004), and haploid viable mutant alleles of yeast *CDC9*, *POB3*, and *GLC7* have been shown to cause sensitivity to sublethal doses of MMS and/or HU (Prakash and Prakash 1977; Schlesinger and Formosa 2000; Bazzi *et al.* 2010). Twelve mutant alleles cause sensitivity to MMS, while two cause increased resistance to MMS. The majority of the mutants sensitive to HU are in h*LIG1* (12/14), which is expected given the major role of that protein in DNA replication. To assess whether computational predictions match empirical results, we used PredictSNP (a consensus classifier grouping six computational tools) (Bendl *et al.* 2014) to predict the effects of the mutations on protein function. Our results demonstrate that predicted effects do not always match observed results. While some substitutions predicted to be neutral with high confidence displayed detrimental growth phenotypes (*e.g.*, h*PPP1CC*-E116D), others predicted to be deleterious with high confidence exhibited no growth defects (*e.g.*, h*SSRP1*-R370C). Moreover, amino acid conservation alone or in combination with computational predictions did not accurately predict observable phenotypes. For instance, an amino acid that is conserved between yeast and human proteins might be expected to be important for protein function. However, even a substitution predicted to be deleterious by computational tools may not result in any growth phenotypes (*e.g.*, h*SSRP1*-T209I). Overall, these results underscore the requirement for direct functional testing of variants in their native gene context.

Discussion

The work presented in this study extends the growing list of human–yeast complementation pairs that will serve as an important resource for model organism and human biology. Many factors impact replaceability of yeast genes by human genes. Different types of yeast strains permit different complementation readouts; therefore, choosing the appropriate yeast strain depends on what is required for downstream applications. We restricted our complementation assays to testing rescue of lethality of the haploid yeast gene knockout because rescue of a yeast conditional allele or downregulated yeast gene can represent partial or indirect complementation. This is evident in the results of Kachroo *et al.* (2015), which indicate that <60% (32/56) of the 69 human orthologs identified by rescue of temperature-sensitive alleles are able to rescue the null allele. Similarly, ∼60% (24/39) of the 44 human orthologs identified by rescue of downregulated yeast genes are able to rescue the null allele. If the main objective is to use yeast as a platform to characterize the functional consequence of human gene variants, we believe that complementation of the null allele is more desirable mainly for two reasons. First, expressing human cDNAs in a yeast deletion mutant diminishes the possibility of misleading phenotypic readouts by eliminating unwanted side effects of the residual yeast protein that is present in partial loss-of-function alleles. Second, the use of null alleles distinguishes those human proteins that are truly substituting for the essential function of the yeast protein from those human proteins that are suppressing a growth phenotype by a secondary mechanism that may or may not relate to the conserved function. Other factors that impact replaceability include the conditions used to express the human cDNA in yeast to account for dosage, toxicity, or timing of expression. Examples of what can be manipulated include picking a suitable promoter (endogenous yeast promoter *vs.* constitutive *vs.* inducible), integration of human cDNA *vs.* episomal vector-based expression, and inclusion of epitope tags. For instance, in this study we used yeast expression vectors that introduce short C-terminal extensions, which can interfere with replaceability by some human proteins (Kachroo *et al.* 2015). In general, there are no sets of conditions that satisfy the replaceability requirements of all candidate complementation pairs, as each human–yeast gene pair is unique. Even when no complementation is observed, partial fusions to create chimeric human–yeast

Table 2 Tumor-specific variants constructed and analyzed in a yeast deletion background

	Nucleotide mutation	Amino acid mutation	Zygosity	Amino acid conserved in yeast	PredictSNP (confidence)	No drug	0.01% MMS	100 mM HU
h*LIG1*	179C > T	A60V	Unknown	No	Neutral (83%)	0.56 ± 0.06***	0.50 ± 0.05	0.38 ± 0.03*
	421A > G	S141G	Heterozygous	Yes	Neutral (75%)	0.45 ± 0.03***	0.50 ± 0.07	0.25 ± 0.05*
	454A > G	K152E	Heterozygous	Yes	Neutral (83%)	0.52 ± 0.06***	0.98 ± 0.06***	0.46 ± 0.03
	455A > G	K152R	Heterozygous	Yes	Neutral (83%)	0.50 ± 0.05***	0.50 ± 0.06	0.23 ± 0.01***
	457G > A	E153K	Heterozygous	No	Neutral (83%)	0.60 ± 0.12***	0.44 ± 0.06	0.30 ± 0.07**
	488G > A	S163N	Unknown	No	Neutral (83%)	0.60 ± 0.12***	0.49 ± 0.05	0.23 ± 0.06***
	664C > T	R222C	Heterozygous	No	Neutral (63%)	0.35 ± 0.01***	0.23 ± 0.01***	0.21 ± 0.02***
	1045G > A	V349M	Unknown	No	Neutral (60%)	0.47 ± 0.03***	0.59 ± 0.04*	0.29 ± 0.05*
	1120G > A	A374T	Unknown	No	Neutral (83%)	0.71 ± 0.03***	0.71 ± 0.06	0.51 ± 0.06**
	1184C > A	P395Q	Unknown	No	Neutral (83%)	0.63 ± 0.05***	0.48 ± 0.04**	0.44 ± 0.04*
	1502T > C	M501T	Unknown	No	Neutral (83%)	0.61 ± 0.06***	0.34 ± 0.01***	0.46 ± 0.03*
	1835C > T	S612L	Heterozygous	No	Neutral (74%)	0.49 ± 0.04***	0.33 ± 0.02***	0.42 ± 0.04
	1969T > G	L657V	Heterozygous	Yes	Neutral (83%)	0.71 ± 0.08***	0.70 ± 0.10	0.44 ± 0.04**
	2290G > A	A764T	Unknown	Yes	Deleterious (87%)	0.73 ± 0.09***	0.30 ± 0.02***	0.61 ± 0.04
	2353G > A	E785K	Unknown	No	Neutral (75%)	0.59 ± 0.07***	0.52 ± 0.05	0.56 ± 0.09
	2446G > A	V816M	Unknown	No	Neutral (74%)	0.70 ± 0.06***	0.36 ± 0.01***	0.47 ± 0.04*
h*SSRP1*	97A > G	K33E	Heterozygous	Yes	Deleterious (87%)	1.02 ± 0.12	0.70 ± 0.08	0.95 ± 0.12
	566C > T	A189V	Heterozygous	Yes	Deleterious (72%)	0.69 ± 0.11**	0.35 ± 0.03	0.79 ± 0.05
	626C > T	T209I	Unknown	Yes	Deleterious (61%)	1.09 ± 0.14	0.68 ± 0.07*	1.00 ± 0.12
	682A > G	K228E	Heterozygous	Yes	Deleterious (87%)	Non-functional	n/d	n/d
	970C > T	R324C	Heterozygous	Yes	Neutral (60%)	1.00 ± 0.06	0.80 ± 0.18	0.97 ± 0.10
	1108C > T	R370C	Heterozygous	No	Deleterious (65%)	1.05 ± 0.11	0.67 ± 0.06*	0.97 ± 0.17
	1306C > T	P436S	Heterozygous	No	Neutral (83%)	0.94 ± 0.05	0.81 ± 0.10	1.01 ± 0.11
	1441T > C	S481P	Unknown	No	Deleterious (61%)	Non-functional	n/d	n/d
	1493A > G	N498S	Heterozygous	No	Neutral (74%)	1.05 ± 0.06	0.70 ± 0.04*	1.08 ± 0.08
	1723A > G	T575A	Heterozygous	No	Neutral (83%)	1.02 ± 0.06	0.76 ± 0.04*	1.22 ± 0.13
	1724C > T	T575M	Unknown	No	Neutral (60%)	1.02 ± 0.05	0.78 ± 0.08*	1.07 ± 0.10
	1950G > T	K650N	Unknown	No	Neutral (83%)	1.07 ± 0.16	0.82 ± 0.03	1.22 ± 0.03
h*PPP1CA*	428G > A	R143H	Heterozygous	Yes	Deleterious (87%)	0.69 ± 0.08	0.72 ± 0.09	0.67 ± 0.03
	815A > G	Y272C	Heterozygous	Yes	Deleterious (87%)	Non-functional	n/d	n/d
h*PPP1CC*	348G > T	E116D	Unknown	Yes	Neutral (75%)	0.57 ± 0.11**	0.24 ± 0.03*	0.33 ± 0.01*
	559C > T	R187W	Unknown	Yes	Deleterious (87%)	0.05 ± 0.01***	0.05 ± 0.01	0.02 ± 0.01*
	601C > T	L201F	Heterozygous	Yes	Deleterious (51%)	1.48 ± 0.18*	1.09 ± 0.24	1.10 ± 0.11
	610C > A	L204I	Unknown	Yes	Neutral (63%)	0.97 ± 0.10	0.89 ± 0.22	1.07 ± 0.14
	866T > G	L289R	Heterozygous	Yes	Deleterious (87%)	Nonfunctional	n/d	n/d

Strain fitness for each allele is expressed as a ratio relative to the yeast strain expressing the wild-type allele grown in the same plate in the same media condition. For "no drug" condition, significance was calculated compared to the wild-type allele grown in the same plate to assess impact of missense mutation on strain fitness. For "drug" condition, significance was calculated compared to the same allele in the no drug condition grown in the same plate to assess impact of drug on strain fitness. Alleles were defined as nonfunctional based on inability to complement the corresponding deletion mutant yeast strain. Strain fitness values that were not determined are indicated (n/d). *$P < 0.01$, **$P < 0.001$, ***$P < 0.0001$.

proteins may allow complementation and provide a useful resource for specific applications (Zhou and Reed 1998).

The experimental design of systematic complementation assays impacts the scope of the results. Testing complementation pairs one-by-one reduces the number of false negatives inherent in pooled screening. This was observed in our study: our one-to-one screen identified 34 complementation pairs corresponding to 28 yeast strains, while our pool-to-pool screen identified 12 of these complementation pairs corresponding to 12 yeast strains, resulting in a 35% recovery rate from the pooled screen (Table 1). This modest recovery rate may be due to several factors: (i) even with 100-fold coverage, shorter cDNAs are shuttled more efficiently from entry clones to destination vectors in Gateway's LR reaction, resulting in a potentially skewed representation of some expression vectors within a pool; (ii) competitive outgrowth can result in a skewed representation of yeast strains pre- and post-transformation; (iii) even with 100-fold coverage, pool-to-pool transformation reduces the chances of a particular human cDNA complementing the matched yeast strain, especially given the possibility of skewed representation of expression vectors and yeast strains; (iv) toxicity from some human cDNAs can skew representation of yeast strains within a pool; (v) sporulation efficiencies of different yeast strains skew haploid selection; and (vi) sequence coverage of rescued haploids will impact the scope of the results. Nevertheless, the advantage of pooled screening is the potential for identification of nonorthologous complementation pairs, as demonstrated in this study. With the development of better screening methods and updated collections, the potential for discovery of additional nonorthologous complementation pairs will aid in deciphering biological mechanisms that would otherwise be overlooked.

We demonstrated a robust approach to determining the functional consequence of human gene variants in yeast.

Because the functional effects of missense mutations are difficult to predict, a quick yet systematic approach to screen and prioritize human gene variants for subsequent testing in mammalian models is of great value. As we have picked mutants of genes implicated in chromosome stability, mutant alleles that result in a nonfunctional protein, or reduction of function as displayed by decreased strain fitness, or increased sensitivity to DNA-damaging agents, are candidate mutations that might contribute to chromosome instability in tumor cells. While we have focused on screening somatic mutations found in tumors, this approach is also applicable to rare-disease variants that are being discovered at an unprecedented rate through next-generation sequencing (Hieter and Boycott 2014). Of course, there are limitations to complementation in model organisms, including representation of multicellular and developmental pathways that are absent in single-celled yeast. Nevertheless, advances in sequencing technologies have led to the discovery of an overwhelming number of human genetic variants whose empirical assessment of their biological effects would be impractical without the support of model organisms.

Acknowledgments

We thank Peter Stirling and Nigel O'Neil for helpful discussion and critical reading of the manuscript. This work was supported by grants to P.H. from the Canadian Institutes of Health Research (CIHR) (MOP 38096) and the National Institutes of Health (R01CA158162). A.H. is supported by doctoral awards from CIHR and the University of British Columbia. C.N. and G.G. are funded by the US Department of Agriculture and the Canada Foundation for Innovation. P.H. is a senior fellow in the Genetic Networks program at the Canadian Institute for Advanced Research.

Literature Cited

Alberti, S., A. D. Gitler, and S. Lindquist, 2007 A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast 24: 913–919.

Arnoldo, A., S. Kittanakom, L. E. Heisler, A. B. Mak, A. I. Shukalyuk *et al.*, 2014 A genome scale overexpression screen to reveal drug activity in human cells. Genome Med. 6: 32.

Balakrishnan, R., J. Park, K. Karra, B. C. Hitz, G. Binkley *et al.*, 2012 YeastMine–an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. Database (Oxford) 2012: bar062.

Barber, T. D., K. McManus, K. W. Yuen, M. Reis, G. Parmigiani *et al.*, 2008 Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. Proc. Natl. Acad. Sci. USA 105: 3443–3448.

Bazzi, M., D. Mantiero, C. Trovesi, G. Lucchini, and M. P. Longhese, 2010 Dephosphorylation of gamma H2A by Glc7/protein phosphatase 1 promotes recovery from inhibition of DNA replication. Mol. Cell. Biol. 30: 131–145.

Bendl, J., J. Stourac, O. Salanda, A. Pavelka, E. D. Wieben *et al.*, 2014 PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLOS Comput. Biol. 10: e1003440.

Beranek, D. T., 1990 Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutat. Res. 231: 11–30.

Brachmann, C. B., A. Davies, G. J. Cost, E. Caputo, J. Li *et al.*, 1998 Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14: 115–132.

Cannon, J. F., 2010 Function of protein phosphatase-1, Glc7, in Saccharomyces cerevisiae. Adv. Appl. Microbiol. 73: 27–59.

Cerami, E., J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer *et al.*, 2012 The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2: 401–404.

Ceulemans, H., and M. Bollen, 2004 Functional diversity of protein phosphatase-1, a cellular economizer and reset button. Physiol. Rev. 84: 1–39.

Charames, G. S., and B. Bapat, 2003 Genomic instability and cancer. Curr. Mol. Med. 3: 589–596.

Dunham, M. J., and D. M. Fowler, 2013 Contemporary, yeast-based approaches to understanding human genetic variation. Curr. Opin. Genet. Dev. 23: 658–664.

Egloff, M. P., P. T. Cohen, P. Reinemer, and D. Barford, 1995 Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. J. Mol. Biol. 254: 942–959.

Egloff, M. P., D. F. Johnson, G. Moorhead, P. T. Cohen, P. Cohen *et al.*, 1997 Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1. EMBO J. 16: 1876–1887.

Forbes, S. A., D. Beare, P. Gunasekaran, K. Leung, N. Bindal *et al.*, 2015 COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 43: D805–D811.

Formosa, T., 2008 FACT and the reorganized nucleosome. Mol. Biosyst. 4: 1085–1093.

Gao, J., B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross *et al.*, 2013 Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6: pl1.

Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles *et al.*, 2002 Functional profiling of the Saccharomyces cerevisiae genome. Nature 418: 387–391.

Greene, A. L., J. R. Snipe, D. A. Gordenin, and M. A. Resnick, 1999 Functional analysis of human FEN1 in Saccharomyces cerevisiae and its role in genome stability. Hum. Mol. Genet. 8: 2263–2273.

Hanahan, D., and R. A. Weinberg, 2011 Hallmarks of cancer: the next generation. Cell 144: 646–674.

Hartley, J. L., G. F. Temple, and M. A. Brasch, 2000 DNA cloning using in vitro site-specific recombination. Genome Res. 10: 1788–1795.

Helenius, J., D. T. Ng, C. L. Marolda, P. Walter, M. A. Valvano *et al.*, 2002 Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. Nature 415: 447–450.

Hieter, P., and K. M. Boycott, 2014 Understanding rare disease pathogenesis: a grand challenge for model organisms. Genetics 198: 443–445.

Hu, Y., A. Rolfs, B. Bhullar, T. V. Murthy, C. Zhu *et al.*, 2007 Approaching a complete repository of sequence-verified protein-encoding clones for Saccharomyces cerevisiae. Genome Res. 17: 536–543.

Johnson, D. I., 1999 Cdc42: An essential Rho-type GTPase controlling eukaryotic cell polarity. Microbiol. Mol. Biol. Rev. 63: 54–105.

Kachroo, A. H., J. M. Laurent, C. M. Yellman, A. G. Meyer, C. O. Wilke *et al.*, 2015 Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science 348: 921–925.

Kato, S., S. Y. Han, W. Liu, K. Otsuka, H. Shibata *et al.*, 2003 Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc. Natl. Acad. Sci. USA 100: 8424–8429.

Kinsella, R. J., A. Kahari, S. Haider, J. Zamora, G. Proctor *et al.*, 2011 Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford) 2011: bar030.

Koc, A., L. J. Wheeler, C. K. Mathews, and G. F. Merrill, 2004 Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. J. Biol. Chem. 279: 223–230.

Koonin, E. V., 2005 Orthologs, paralogs, and evolutionary genomics. Annu. Rev. Genet. 39: 309–338.

Lengauer, C., K. W. Kinzler, and B. Vogelstein, 1997 Genetic instability in colorectal cancers. Nature 386: 623–627.

Lindahl, T., and D. E. Barnes, 1992 Mammalian DNA ligases. Annu. Rev. Biochem. 61: 251–281.

Marini, N. J., P. D. Thomas, and J. Rine, 2010 The use of orthologous sequences to predict the impact of amino acid substitutions on protein function. PLoS Genet. 6: e1000968.

Mayfield, J. A., M. W. Davies, D. Dimster-Denk, N. Pleskac, S. McCarthy *et al.*, 2012 Surrogate genetics and metabolic profiling for characterization of human disease alleles. Genetics 190: 1309–1323.

McCarrey, J. R., and K. Thomas, 1987 Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. Nature 326: 501–505.

Osborn, M. J., and J. R. Miller, 2007 Rescuing yeast mutants with human genes. Brief. Funct. Genomics Proteomics 6: 104–111.

Osborne, A. R., T. A. Rapoport, and B. van den Berg, 2005 Protein translocation by the Sec61/SecY channel. Annu. Rev. Cell Dev. Biol. 21: 529–550.

Pan, X., D. S. Yuan, D. Xiang, X. Wang, S. Sookhai-Mahadeo *et al.*, 2004 A robust toolkit for functional profiling of the yeast genome. Mol. Cell 16: 487–496.

Pan, X., D. S. Yuan, S. L. Ooi, X. Wang, S. Sookhai-Mahadeo *et al.*, 2007 dSLAM analysis of genome-wide genetic interactions in Saccharomyces cerevisiae. Methods 41: 206–221.

Prakash, L., and S. Prakash, 1977 Isolation and characterization of MMS-sensitive mutants of Saccharomyces cerevisiae. Genetics 86: 33–55.

Schlesinger, M. B., and T. Formosa, 2000 POB3 is required for both transcription and replication in the yeast *Saccharomyces cerevisiae*. Genetics 155: 1593–1606.

Schvartzman, J. M., R. Sotillo, and R. Benezra, 2010 Mitotic chromosomal instability and cancer: mouse modelling of the human disease. Nat. Rev. Cancer 10: 102–115.

Stirling, P. C., M. S. Bloom, T. Solanki-Patil, S. Smith, P. Sipahimalani *et al.*, 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.

Tarnowski, L. J., P. Kowalec, M. Milewski, M. Jurek, D. Plochocka *et al.*, 2012 Nuclear import and export signals of human cohesins SA1/STAG1 and SA2/STAG2 expressed in Saccharomyces cerevisiae. PLoS One 7: e38740.

Tugendreich, S., E. Perkins, J. Couto, P. Barthmaier, D. Sun *et al.*, 2001 A streamlined process to phenotypically profile heterologous cDNAs in parallel using yeast cell-based assays. Genome Res. 11: 1899–1912.

VanDemark, A. P., M. Blanksma, E. Ferris, A. Heroux, C. P. Hill *et al.*, 2006 The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. Mol. Cell 22: 363–374.

Yang, X., J. S. Boehm, X. Yang, K. Salehi-Ashtiani, T. Hao *et al.*, 2011 A public genome-scale lentiviral expression library of human ORFs. Nat. Methods 8: 659–661.

Yuen, K. W., C. D. Warren, O. Chen, T. Kwok, P. Hieter *et al.*, 2007 Systematic genome instability screens in yeast and their potential relevance to cancer. Proc. Natl. Acad. Sci. USA 104: 3925–3930.

Zhang, L., Z. Zhang, F. Long, and E. Y. Lee, 1996 Tyrosine-272 is involved in the inhibition of protein phosphatase-1 by multiple toxins. Biochemistry 35: 1606–1611.

Zhang, N., M. Osborn, P. Gitsham, K. Yen, J. R. Miller *et al.*, 2003 Using yeast to place human genes in functional categories. Gene 303: 121–129.

Zhou, Z., and R. Reed, 1998 Human homologs of yeast prp16 and prp17 reveal conservation of the mechanism for catalytic step II of pre-mRNA splicing. EMBO J. 17: 2095–2106.

Communicating editor: M. Johnston

GENETICS

Supporting Information
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.181099/-/DC1

Complementation of Yeast Genes with Human Genes as an Experimental Platform for Functional Testing of Human Genetic Variants

Akil Hamza, Erik Tammpere, Megan Kofoed, Christelle Keong, Jennifer Chiang, Guri Giaever, Corey Nislow, and Philip Hieter

DOI: 10.1534/genetics.115.181099

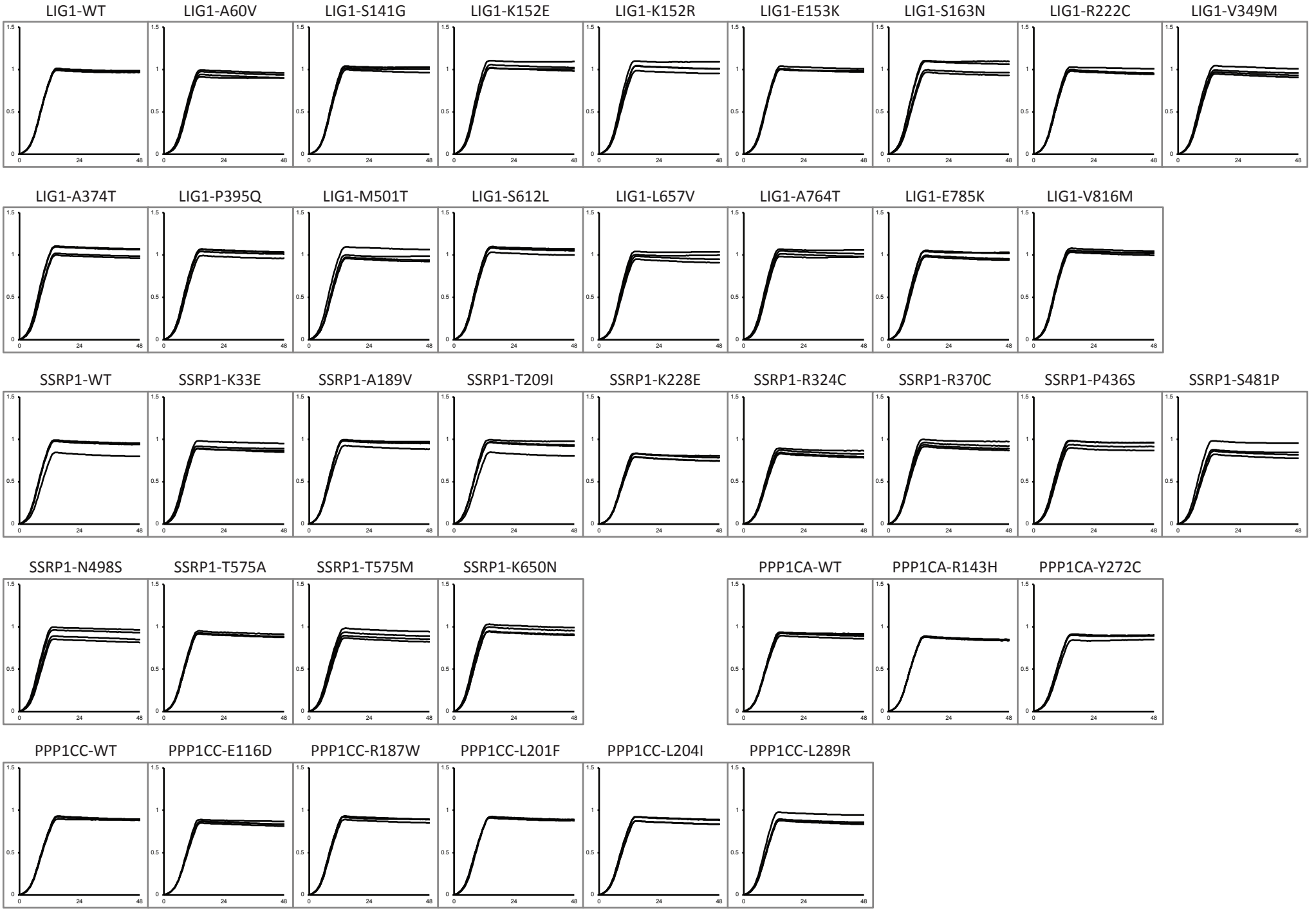

Figure S1. Growth curve replicates of wild-type yeast (BY4742) expressing indicated human cDNAs in –Leu media at 30°C. X-axis represents time in hours, while y-axis represents OD600 readings.

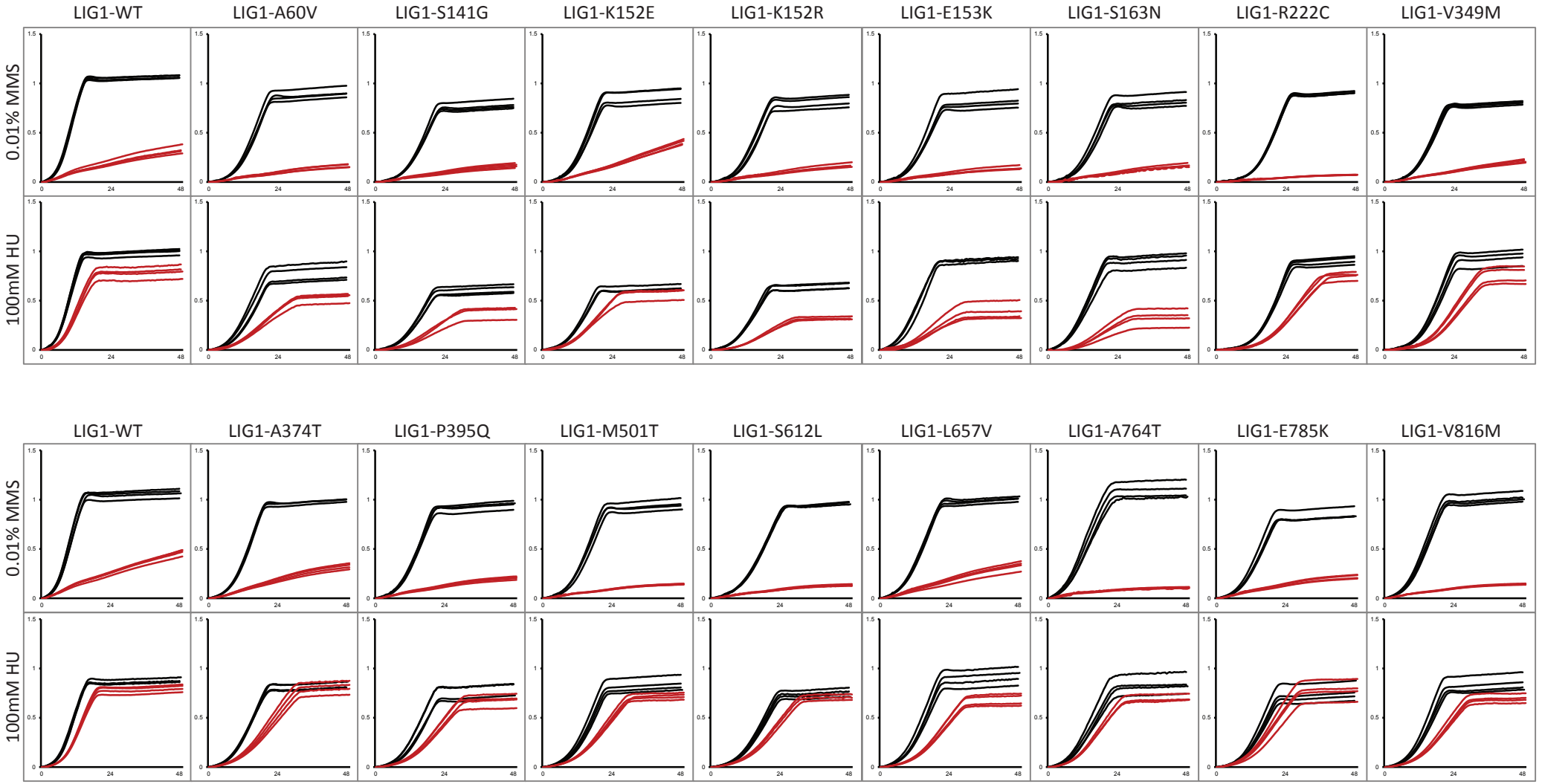

Figure S2. Growth curve replicates of *cdc9*Δ yeast strains expressing indicated human cDNAs in –Leu media at 30°C. Black lines represent growth curve replicates in "No Drug" condition, while red lines represent growth curve replicates in indicated "Drug" condition. The "No Drug" and "Drug" replicates shown in each panel were grown on the same plate and diluted from the same mid-log phase culture. X-axis represents time in hours, while y-axis represents OD600 readings.

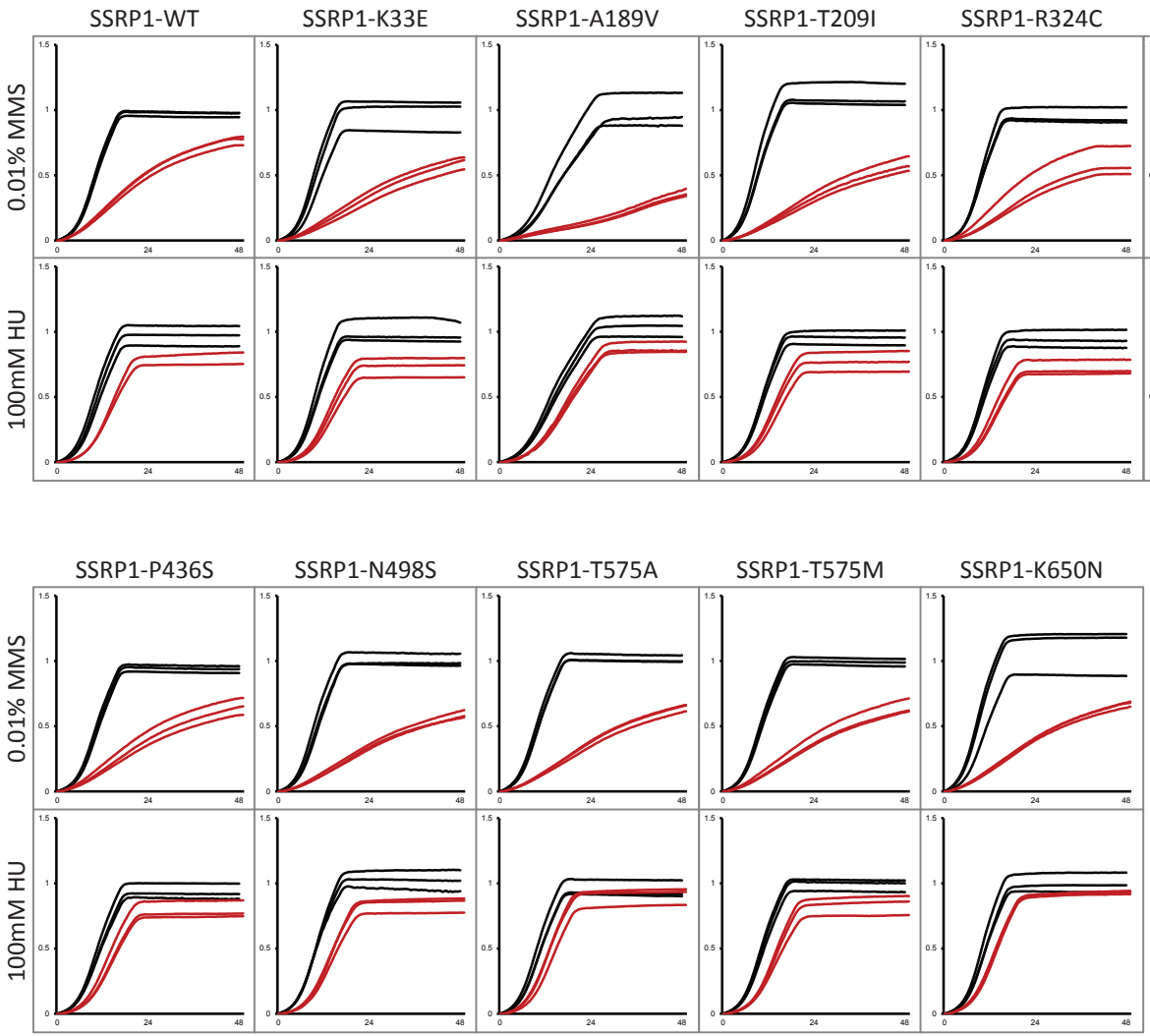

Figure S3. Growth curve replicates of *pob3*Δ yeast strains expressing indicated human cDNAs in –Leu media at 30°C. Black lines represent growth curve replicates in "No Drug" condition, while red lines represent growth curve replicates in indicated "Drug" condition. The "No Drug" and "Drug" replicates shown in each panel were grown on the same plate and diluted from the same mid-log phase culture. X-axis represents time in hours, while y-axis represents OD600 readings.

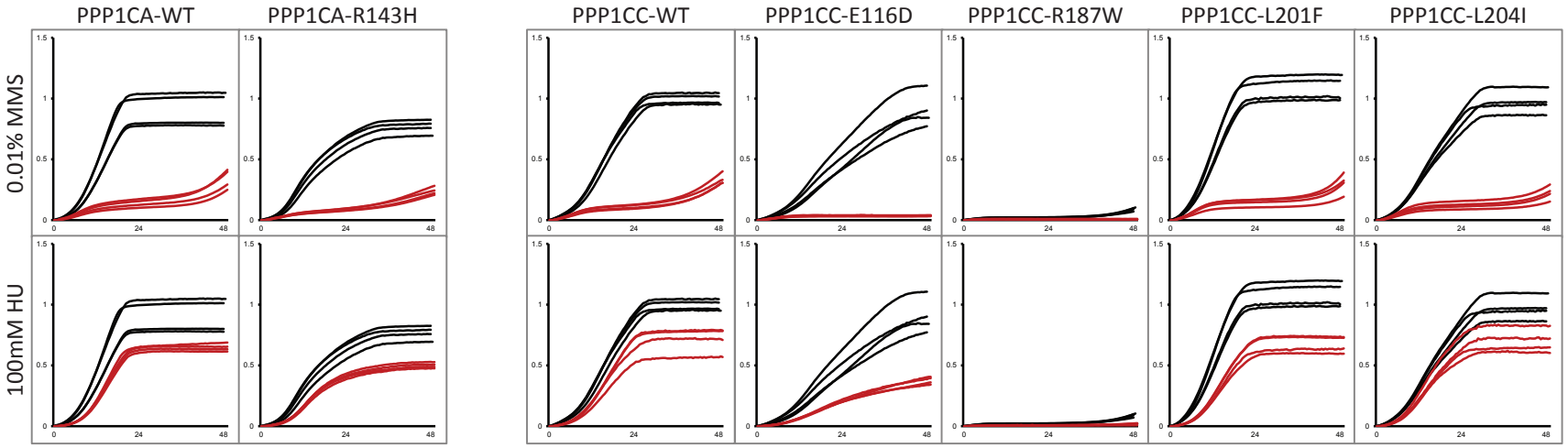

Figure S4. Growth curve replicates of *glc7*Δ yeast strains expressing indicated human cDNAs in –Leu media at 30°C. Black lines represent growth curve replicates in "No Drug" condition, while red lines represent growth curve replicates in indicated "Drug" condition. The "No Drug" and "Drug" replicates shown in each panel were grown on the same plate and diluted from the same mid-log phase culture. X-axis represents time in hours, while y-axis represents OD600 readings.

Tables S1-S6
Available for download as Excel files at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.181099/-/DC1

Table S1. Plasmids and strains used for screening tumor-specific variants in yeast

Table S2. Essential yeast CIN genes and human homologs tested in the one-to-one complementation screen

Table S3. Essential yeast genes and human homologs included in the pool-to-pool complementation screen

Table S4. Comparison of our compiled list of complementation pairs to literature sources

Table S5. Select examples of essential yeast genes from the one-to-one screen that had multiple homologs tested for complementation

Table S6. Tumor-specific variants analyzed in a yeast wild-type background