An integrated platform of genomic assays reveals small-molecule bioactivities

Shawn Hoon[1,2], Andrew M Smith[3–5], Iain M Wallace[4], Sundari Suresh[2], Molly Miranda[2], Eula Fung[2], Michael Proctor[2], Kevan M Shokat[6], Chao Zhang[6], Ronald W Davis[1,2], Guri Giaever[3,5,7], Robert P St Onge[2] & Corey Nislow[3–5]

Bioactive compounds are widely used to modulate protein function and can serve as important leads for drug development. Identifying the *in vivo* targets of these compounds remains a challenge. Using yeast, we integrated three genome-wide gene-dosage assays to measure the effect of small molecules *in vivo*. A single TAG microarray was used to resolve the fitness of strains derived from pools of (i) homozygous deletion mutants, (ii) heterozygous deletion mutants and (iii) genomic library transformants. We demonstrated, with eight diverse reference compounds, that integration of these three chemogenomic profiles improves the sensitivity and specificity of small-molecule target identification. We further dissected the mechanism of action of two protein phosphatase inhibitors and in the process developed a framework for the rational design of multidrug combinations to sensitize cells with specific genotypes more effectively. Finally, we applied this platform to 188 novel synthetic chemical compounds and identified both potential targets and structure-activity relationships.

Identifying the mechanism of action of bioactive compounds *in vivo* is essential for chemical biology and drug discovery. To realize this goal, a number of approaches using cell-based assays to carry out forward chemical genetic screens in yeast have been developed. For example, drug-induced haploinsufficiency profiling (HIP) was developed to identify small molecules that target essential genes[1,2]. Similarly, homozygous profiling (HOP) uses the yeast homozygous (or haploid) deletion collection to identify genetic modifiers of drug resistance[3,4]. Diverse multicopy suppression strategies have been described that identify genes that confer resistance to chemical treatment when overexpressed[5–7]. Other strategies rely on 'guilt by association'— inferring targets from a compendium of reference profiles (for example, genetic interactions and gene expression)[8–11].

Each method has limitations, and here we present a cost-effective method that integrates three distinct assays with the goal of improving the characterization of the bioactivity of small molecules and making this technology accessible. Using a miniaturized screening procedure in conjunction with a common barcode TAG microarray[12,13], this platform systematically examines the effect of both increasing and decreasing gene dosage on chemical sensitivity. We demonstrate its utility for small-molecule characterization with eight well-studied chemical compounds with distinct mechanisms of action (**Supplementary Table 1** online). We successfully identified both known and new molecular targets for several of these compounds and noted that

examining the effect of both increasing and decreasing gene dosage was often required to distinguish the bona fide target from a longer list of potential candidates. This multipronged approach, combined with genome-wide drug combination screens, uncovered important *in vivo* differences between the phosphatase inhibitors cantharidin and calyculin A. We also identified a previously uncharacterized gene, which we dub *CRG1* (cantharidin resistance gene 1), as a dose-dependent regulator of resistance to cantharidin, a natural product isolated from blister beetles. Extension of multicopy suppression profiling identified a potential *CRG1* ortholog in the fungal pathogen *Candida albicans* (*orf19.633*), which underscores the flexibility of this assay and the utility of small molecules for annotating gene function. We also used this platform to identify and characterize drug-drug interactions. Finally, we applied this platform to 188 synthetic chemical compounds with no previously known biological activity. This discovery effort uncovered several potential protein targets and revealed new structure-activity relationships for these compounds.

RESULTS
A chemogenomics platform
The yeast deletion collection is a proven resource for chemical genetics[14]. Each strain in this collection contains a unique 20-base-pair DNA tag that enables the fitness of individual strains (from a heterogeneous pool) to be resolved simultaneously using an

[1]Department of Genetics, Stanford University, Mail Stop-5120, Palo Alto, California 94305, USA. [2]Stanford Genome Technology Center, 855 California Avenue, Palo Alto, California 94304, USA. [3]Department of Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S1A8, Canada. [4]Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ontario M5S3E1, Canada. [5]Donnelly Centre for Cellular and Biomedical Research, University of Toronto, 160 College Street, Toronto, Ontario M5S3E1, Canada. [6]Department of Molecular Pharmacology, University of California, 600 16th Street, MC 2280, San Francisco, California 94158-2280, USA. [7]Department of Pharmaceutical Sciences, University of Toronto, 144 College Street, Toronto, Ontario M5S3M2, Canada. Correspondence should be addressed to R.P.S. (bstonge@stanford.edu) or C.N. (corey.nislow@utoronto.ca).

Figure 1 An integrated chemogenomics screening platform. (**a**) Illustration of the chemogenomic platform that interrogates three different yeast pools with a single TAG4 array. A homozygous deletion pool ($n = 4,990$), a heterozygous deletion pool ($n = 1,145$) and a pool of genomic library transformants ($n \sim 4,700$) are each challenged with a compound of interest (X). The heterozygous deletion pool represents only those genes that are essential for viability or for which homozygous deletions could not be constructed systematically (see Methods). Barcode sequences are isolated and amplified from deletion strain pools, and genomic DNA inserts are isolated and amplified from the library transformants. Labeled products are then hybridized sequentially to the same TAG4 array. (**b**) Genes directly related to a drug's mechanism of action are predicted to alter drug resistance when deleted or overexpressed. Genes that confer both resistance when overexpressed and sensitivity when deleted are more likely to be directly related to the drug's mechanism of action.
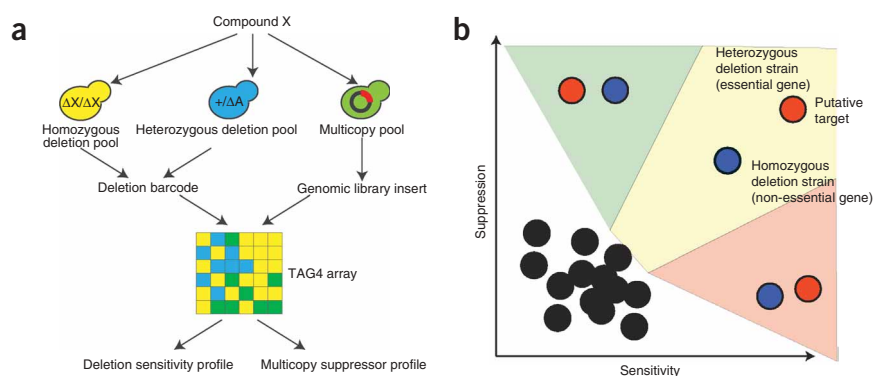
oligonucleotide array. Using a pooled strategy has a distinct advantage over strain-by-strain experiments because it substantially reduces compound usage and simplifies sample processing. Exploiting an array that contains both tag and open reading frame (ORF) probes[12], we developed a chemogenomics platform to resolve the fitness of strains derived from pools of (i) homozygous deletion mutants, (ii) heterozygous deletion mutants and (iii) genomic library transformants (**Fig. 1a**). We refer to those screens that use the deletion pools as deletion sensitivity profiling (DSP) and screens that use genomic library transformants as multicopy suppression profiling (MSP). MSP is a new genome-wide assay that screens DNA clone libraries competitively to identify genes that confer resistance to compounds when overrepresented. Traditional multicopy suppressor screens involve cumbersome plating techniques and clone characterization[5], but more recently, a microarray-based approach was used to characterize small-molecule modulators of the target of rapamycin (TOR) pathway[7]. Here we apply a variation of this concept by using a high-copy, random genomic library (rather than an inducible ORF library)

and a simplified means to amplify library clones (**Supplementary Fig. 1a,b** online). We miniaturized and then validated this assay using a mutant *CDC28* that is specifically sensitive to inhibition by the ATP analog 1-NM-PP1 (**1**) (see **Supplementary Fig. 1c–e** and **Supplementary Methods** online).

Integrating gene-dose assays improves target identification

Several studies have demonstrated that increasing or decreasing the abundance of a small molecule's target will directly affect the sensitivity of the cell to that small molecule[4–7,15,16]. We applied both DSP and MSP to eight reference compounds that have distinct mechanisms of action (**Supplementary Table 1**). The results are presented such that each ORF is represented as a vector in a Cartesian plane whereby the y coordinate represents the level of multicopy suppression and the x coordinate represents the level of deletion sensitivity in response to chemical exposure (**Fig. 1b**). Vector analysis[17] was used to identify strains (colored red or blue) that were significant in both assays (see Methods). The plots are divided into eight different sectors to
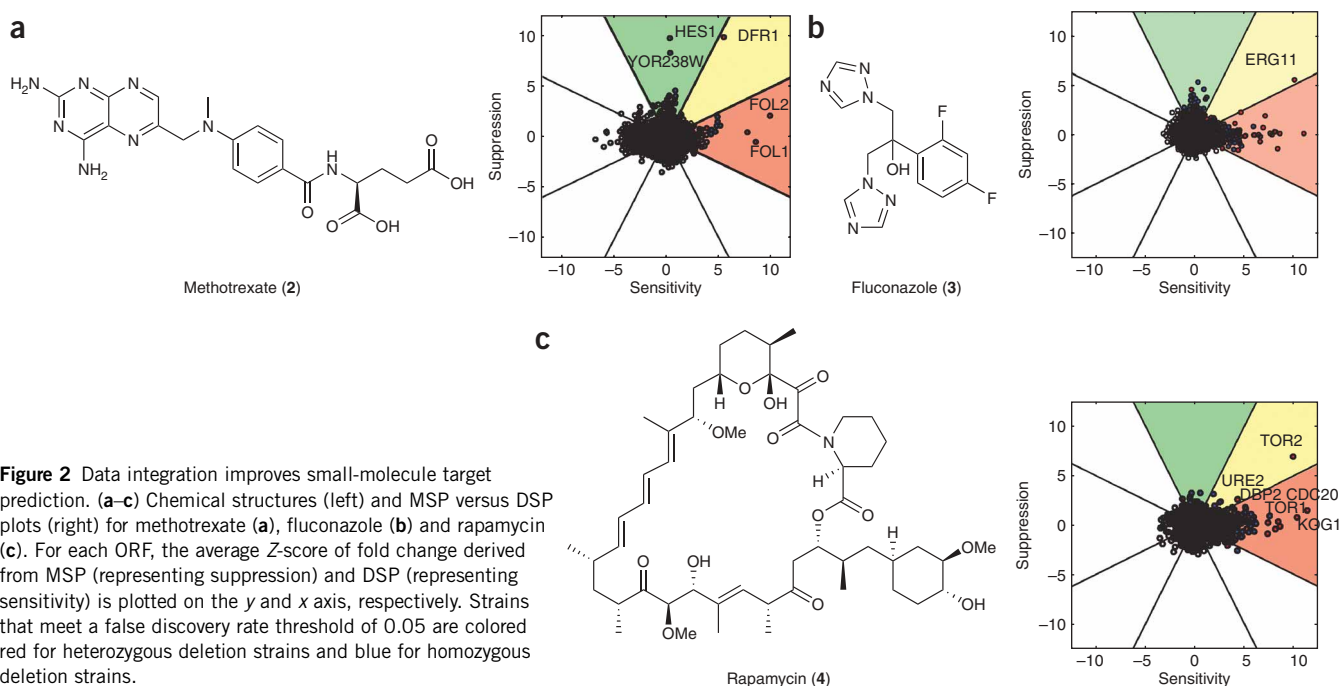
Figure 2 Data integration improves small-molecule target prediction. (**a–c**) Chemical structures (left) and MSP versus DSP plots (right) for methotrexate (**a**), fluconazole (**b**) and rapamycin (**c**). For each ORF, the average Z-score of fold change derived from MSP (representing suppression) and DSP (representing sensitivity) is plotted on the y and x axis, respectively. Strains that meet a false discovery rate threshold of 0.05 are colored red for heterozygous deletion strains and blue for homozygous deletion strains.

represent different response patterns; in this study, we focused on three sectors in particular. The green sector identifies genes that are either bona fide suppressors or genes that are linked to these suppressors (genomic clones often contain multiple genes) but that do not cause sensitivity when deleted. The red sector identifies genes that are sensitive to the chemical when deleted (homozygous or heterozygous) but are not multicopy suppressors. The yellow sector contains genes that result in sensitivity when deleted and result in suppression when overexpressed. Essential genes (red) in this sector represent potential targets, whereas non-essential genes (blue) represent genetic modifiers of drug resistance.

Methotrexate (2), fluconazole (3) and rapamycin (4) are widely prescribed drugs with well-characterized targets and complex mechanisms of action. Methotrexate inhibits folic acid biosynthesis by targeting dihydrofolate reductase (encoded by *DFR1*)[18]. Fluconazole inhibits ergosterol biosynthesis by inhibiting the cytochrome P450 lanosterol 14α-demethylase encoded by *ERG11* (ref. 19). Rapamycin acts by binding to the immunophilin *FKBP12*, and together, the *FKBP12*–rapamycin complex binds to and inhibits the TOR proteins[20]. Consistent with these studies, our experiments identified *DFR1*, *ERG11* and *TOR2* as both haploinsufficient and multicopy suppressors of methotrexate, fluconazole and rapamycin, respectively (**Fig. 2**). We note that in all cases, integrating deletion sensitivity with multicopy suppression greatly refined the list of potential targets. For example, in each case, the most sensitive deletion strain determined by DSP did not represent the respective target (although the bona fide target was always one of the top ten most sensitive strains). Consistent with previous reports, an *FPR1* heterozygous mutant was moderately

resistant to rapamycin (see Methods), whereas *FPR1* overexpression did not modify rapamycin sensitivity[7,21]. Genome-wide results for methylmethane sulfonate (MMS, **5**), latrunculin A (**6**) and nocodazole (**7**) are presented in **Supplementary Figure 2a–c** online.

Surprisingly, we found by sequencing that the major DNA fragment obtained following an MSP screen contained not the full-length *TOR2* gene but rather only the C-terminal portion of the genes *EAP1* and *TOR2* (**Supplementary Fig. 2d**). Because this C-terminal fragment of the *TOR2* gene included the *FKB12*–rapamycin binding domain (FRB), this result implied that overexpression of this domain alone is sufficient to confer rapamycin resistance. We tested this hypothesis directly by expressing a fragment of *TOR2* lacking the FRB domain and showed that loss of the FRB domain abolishes rapamycin resistance (**Supplementary Fig. 2d**). Though this observation is not surprising, it underscores a benefit of performing MSP with DNA fragments—that is, it is useful for defining the ligand binding sites of proteins.

Cantharidin and calyculin A have distinct effects *in vivo*

Cantharidin (**8**) and calyculin A (**9**) are naturally occurring toxins that each inhibit both type I (PP1) and type IIA (PP2A) protein phosphatases[22–24]. These inhibitors have modest *in vitro* specificity for each protein, with cantharidin being more selective for PP2A (ref. 25). We screened both compounds using our integrated platform and found that Glc7, the only essential type I protein phosphatase in yeast, is haploinsufficient in both cantharidin and calyculin A. Notably, overexpression of *GLC7* conferred resistance to calyculin A but not cantharidin (**Fig. 3a,b**). These results were confirmed using isogenic
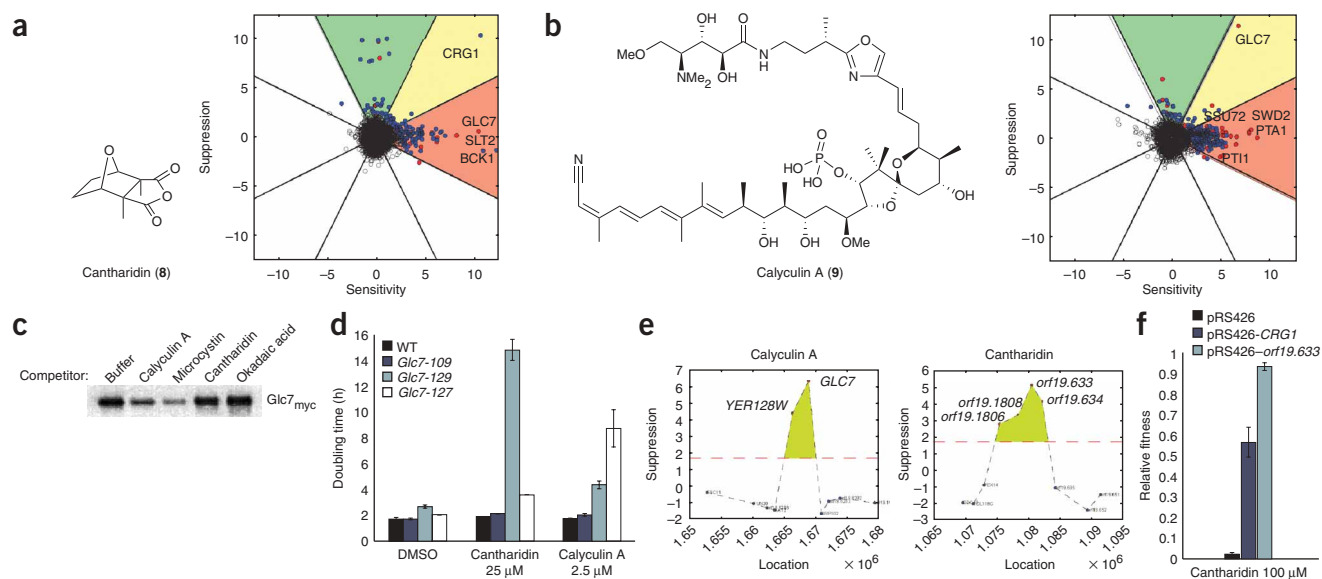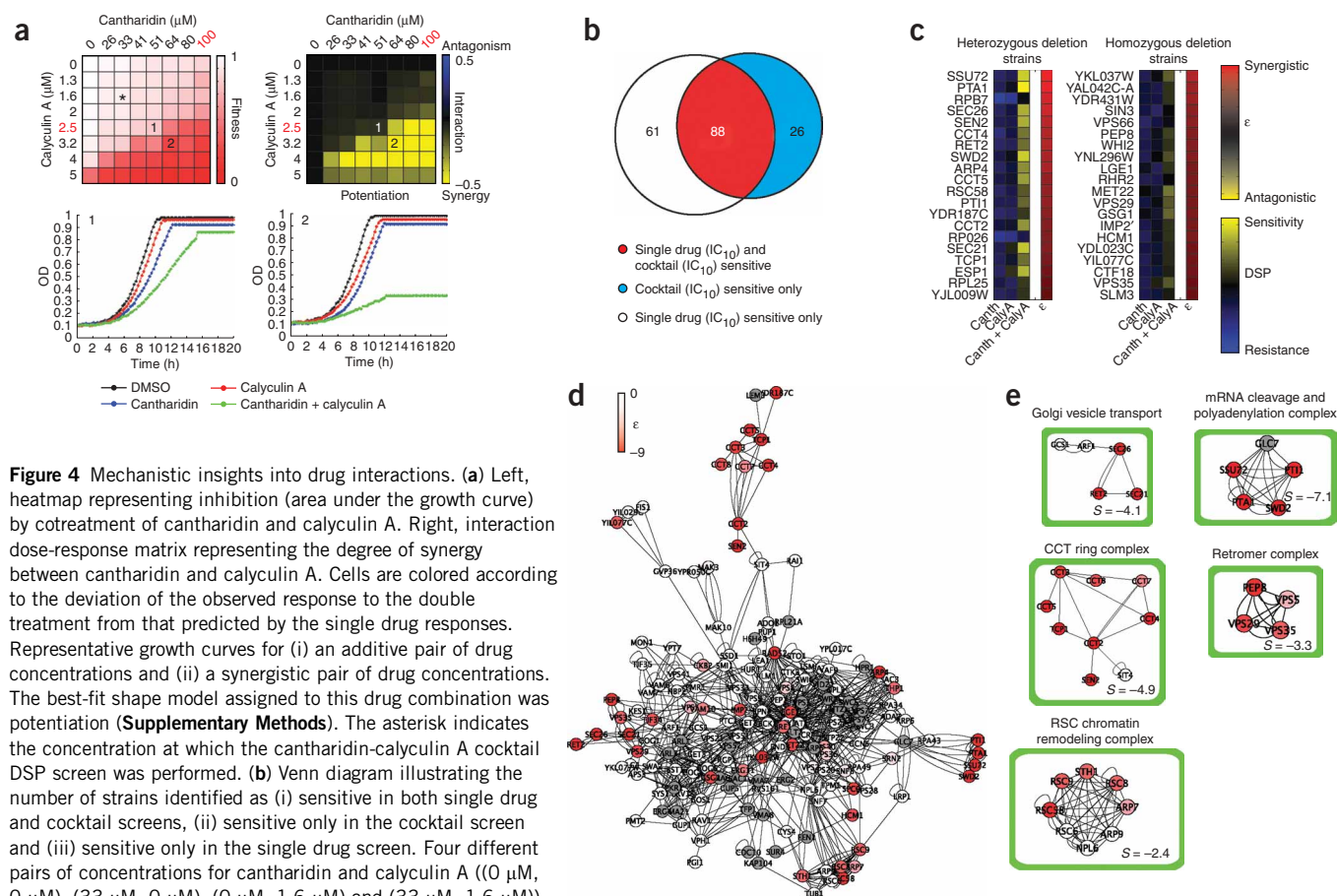

Figure 3 Cantharidin and calyculin A have distinct effects *in vivo*. (**a,b**) Chemical structures (left) and MSP versus DSP plots (right) for cantharidin (**a**) and calyculin A (**b**). (**c**) The relative affinity of calyculin A and cantharidin for Glc7 was measured using a microcystin competition assay. Cell lysates from yeast expressing Glc7$_{myc}$ were pre-incubated with buffer, calyculin A, microcystin, cantharidin or okadaic acid before adding microcystin-conjugated agarose beads. Proteins bound to the beads were isolated, separated by SDS-PAGE and immunoblotted with an anti-myc mouse monoclonal antibody and horseradish peroxidase (HRP)-labeled goat anti-mouse secondary antibody. (**d**) Three different *glc7* alleles, defective in nuclear (*glc7-127*, *glc7-129*) and cytoplasmic (*glc7-109*) pathways, were grown in the presence of cantharidin and calyculin A, and their doubling times are presented ± s.d. (*n* = 3). (**e**) A *C. albicans* genomic library was transformed into *S. cerevisiae* (BY4743). Transformants were pooled and grown in the presence of calyculin A and cantharidin. Recovered inserts were hybridized to a *C. albicans* expression array. The fold ratio (log$_2$ (treatment/control)) of each ORF is plotted on the *y* axis and arranged by its annotated genomic location on the *x* axis. The red dashed line is the fold-change cutoff (log$_2$ 1.6) used to identify each significant suppressor locus. Area highlighted in green corresponds to the identified suppressor locus. (**f**) *orf19.633* and *CRG1* were cloned into the high-copy plasmid pRS426, transformed into BY4743 and grown in the presence and absence of cantharidin (100 μM). Doubling times were normalized to growth of each strain in DMSO control ± s.d. (*n* = 3).

cultures (**Supplementary Fig. 3a** online). Next, we used an *in vitro* assay to measure direct binding of each inhibitor to Glc7 by testing the ability of each inhibitor to disrupt the interaction between Glc7 and microcystin-LR (**10**), a non-cell-permeable inhibitor of Glc7 (ref. 26). These results revealed that calyculin A, but not cantharidin, is an effective competitor for Glc7 binding (**Fig. 3c**). Collectively, these results suggest that calyculin A is a direct inhibitor of the yeast Glc7 protein *in vivo*.

Cantharidin and calyculin A represent the most mechanistically similar compounds in our reference set (**Supplementary Table 1**). These *in vitro* data are not, however, supported by hierarchical clustering of the deletion sensitivity profiles, which revealed that the cantharidin and calyculin A profiles are among the most dissimilar in our dataset (**Supplementary Fig. 3b**). The difference between the two profiles is also reflected by the functional enrichment of strains that are sensitive to each compound (**Supplementary Fig. 3c**). These results underscore differences between the *in vivo* effects of these two phosphatase inhibitors. Further inspection of their deletion sensitivity profiles revealed that each compound identified different *GLC7*-related genes. For example, strains heterozygous for components of the APT (associated with Pta1) complex (*PTA1*, *SWD2*), an essential subcomplex of the yeast cleavage and polyadenylation factor (CPF)-containing complex (holo-CPF)[27,28] that includes Glc7, were

sensitive to both compounds. In contrast, *GLC7* is also known to promote cell wall integrity, bud morphology and polarization of the actin cytoskeleton through the phosphokinase C (PKC)-regulated *SLT2* mitogen-activated protein (MAP) kinase pathway[29], and strains deficient in various components of PKC signaling (*BCK1*, *SLT2*, *RLM1*) were uniquely sensitive to cantharidin. In an independent confirmation of the observed genome-wide differences, we found that three *glc7* alleles, each known to affect different *GLC7*-dependent processes[30–32], were differentially sensitive to each inhibitor (**Fig. 3d**). *glc7* alleles defective in nuclear pathways (*glc7-127*, *glc7-129*) were more sensitive to either compound compared with a *glc7* allele that is defective in cytoplasmic function (*glc7-109*) (**Fig. 3d**). This allele specificity is consistent with the enrichment of nuclear functions (mRNA processing and chromatin modification) that we observed in the genome-wide deletion sensitivity screens. More importantly, *glc7-129* was more sensitive to cantharidin than *glc7-127*, whereas the opposite was true for calyculin A (**Fig. 3d**).

DSP of cantharidin showed that homozygous deletions of *YHR209W* (*CRG1*), a non-essential gene encoding a putative *S*-adenosylmethionine (SAM)-dependent methyltransferase of unknown function, sensitized yeast to cantharidin (**Fig. 3a**). This gene was also haploinsufficient in cantharidin[33] and was an effective multicopy suppressor of cantharidin sensitivity (**Fig. 3a** and

Figure 4 Mechanistic insights into drug interactions. (**a**) Left, heatmap representing inhibition (area under the growth curve) by cotreatment of cantharidin and calyculin A. Right, interaction dose-response matrix representing the degree of synergy between cantharidin and calyculin A. Cells are colored according to the deviation of the observed response to the double treatment from that predicted by the single drug responses. Representative growth curves for (i) an additive pair of drug concentrations and (ii) a synergistic pair of drug concentrations. The best-fit shape model assigned to this drug combination was potentiation (**Supplementary Methods**). The asterisk indicates the concentration at which the cantharidin-calyculin A cocktail DSP screen was performed. (**b**) Venn diagram illustrating the number of strains identified as (i) sensitive in both single drug and cocktail screens, (ii) sensitive only in the cocktail screen and (iii) sensitive only in the single drug screen. Four different pairs of concentrations for cantharidin and calyculin A ((0 μM, 0 μM), (33 μM, 0 μM), (0 μM, 1.6 μM) and (33 μM, 1.6 μM)) were screened using DSP. We classified a strain as sensitive to the single drug treatment if it was sensitive to either cantharidin (100 μM) or calyculin A (2.5 μM). (**c**) Heatmap representing strain sensitivity (blue to yellow) and drug interaction (yellow to red) for the top 20 deletion strains exhibiting synergistic interactions. (**d**) An interaction network with nodes colored (white to red) by degree of drug synergy as quantified by ε. Gray nodes represent strains for which ε values could not be accurately assigned (see Methods). (**e**) Modules representing protein complexes with the highest average synergy score *S*.

Table 1 20 gene-compound pairs with the highest combined DSP-MSP activity scores

|	Strain	Compound	Activity score	Strain
1	*CRG1*	Cantharidin (**8**)	2,082.2	hom
2	*TOR2*	Rapamycin (**4**)	1,106.9	het
3	*ERG11*	Fluconazole (**3**)	1,089.1	het
4	*SEC14*	4130-1278 (**11**)	1,011.7	het
5	*ERG11*	4513-0042 (**13**)	702.6	het
6	*GLC7*	Calyculin A (**9**)	527.9	het
7	*ERG24*	4092-0821 (**18**)	463.4	hom
8	*PDC1*	0086-0128 (**19**)	354.4	hom
9	*SOD2*	3013-0144 (**20**)	324.9	hom
10	*HEM1*	0986-0246 (**21**)	316.5	het
11	*PDR1*	3937-0236 (**22**)	314.4	hom
12	*TRP2*	1486-1328 (**23**)	306.5	hom
13	*DFR1*	Methotrexate (**2**)	280.3	het
14	*ERG3*	0958-0271 (**24**)	212.1	hom
15	*HEM15*	1120-0019 (**25**)	201.5	het
16	*YAP1*	0987-0079 (**26**)	179.7	hom
17	*RPN4*	0988-0037 (**27**)	173.5	hom
18	*HYP2*	4466-0038 (**28**)	168.4	het
19	*STP1*	0109-0045 (**29**)	161.6	hom
20	*PAB1*	1326-1318 (**30**)	152.5	het

Supplementary Fig. 4 online). Homology comparisons revealed that *CRG1*, like other SAM-dependent methyltransferases, is poorly conserved outside of a highly conserved structural fold[34]. To identify a functional homolog of *CRG1*, we modified MSP by introducing a *C. albicans* genomic library into *Saccharomyces cerevisiae* to generate a pool of multicopy transformants. This pool was grown in the presence of calyculin A and cantharidin, and resistant clones were identified by hybridization to *C. albicans* expression arrays. Consistent with our screens using a *S. cerevisiae* library, the *C. albicans* ortholog of *GLC7* (*orf19.6285*) was identified as the major suppressor of calyculin A–induced sensitivity (**Fig. 3e**). The predominant suppressor clone identified in the presence of cantharidin contained four genes of unknown function (**Fig. 3e**). The highest scoring gene in this clone (and indeed the entire experiment) was *orf19.633*, annotated as a putative methyltransferase with no ortholog in *S. cerevisiae* based on sequence similarity[35]. Nonetheless, the methyltransferase domain shared between *CRG1* and *orf19.633* suggests that they are functional homologs, and indeed a clone that contained only *orf19.633* conferred even greater cantharidin resistance than a clone containing *S. cerevisiae CRG1* (**Fig. 3f**).

Mechanistic insights into drug interactions

Because cantharidin and calyculin A disrupt distinct but related aspects of cell physiology, we examined the effect of treating cells with both compounds simultaneously. We quantified the degree of interaction across a dose-response matrix surface using a Bliss interaction model[36] and found that the

combination of cantharidin and calyculin A was highly synergistic compared with the combination of cantharidin and seven other compounds (**Fig. 4a** and **Supplementary Fig. 5** online). Of four previously described shape models[37], potentiation, which describes a combination where one drug's curve is shifted with a power-law slope p above an enhancer concentration Y_{POT} (ref. 37), best described the relationship between cantharidin and calyculin A (**Fig. 4a**). In this case, cantharidin strongly potentiated calyculin A with $P = 0.84$ and $Y_{pot} = 41$ μM (**Supplementary Methods**).

To identify gene deletions that specifically sensitized cells to this drug cocktail, DSP was performed in diluent only (DMSO), cantharidin, calyculin A and a cocktail of cantharidin and calyculin A. At the same significance cutoff (false discovery rate < 0.1), fewer strains were sensitive to the cocktail ($n = 114$) compared with the combined number of strains sensitive to either compound at the equivalent inhibitory concentration ($n = 149$) (**Fig. 4b** and **Supplementary Fig. 6a** online). Notably, the cocktail screen also identified additional strains that were not identified by either single-drug treatment ($n = 26$). We next calculated the deviation of the observed fitness from the expectation (of the Bliss model), ε (where $ε = W_{XY} − W_X \cdot W_Y$), for each deletion strain using the array results for single- and double-drug treatments (see Methods). Consistent with the effect on wild type at the selected concentrations, the distribution of ε values is centered at zero, which indicates that the effects of each drug combine predictably for most deletion strains (**Supplementary Fig. 6b**). The 20 heterozygous and 20 homozygous deletion strains with the smallest (most negative) ε scores are listed in **Figure 4c**. These represent gene deletions that sensitize the cell to the cantharidin–calyculin A cocktail. We found that the top 1% of sensitized strains ($n = 56$) were significantly enriched for strains that are sensitive to calyculin A alone, compared with cantharidin alone (chi-square $P < 6.3 \times 10^{-10}$; **Supplementary Fig. 6c**). In agreement with our dose-response surface analysis, this suggests that cantharidin acts by potentiating the effect of calyculin A.
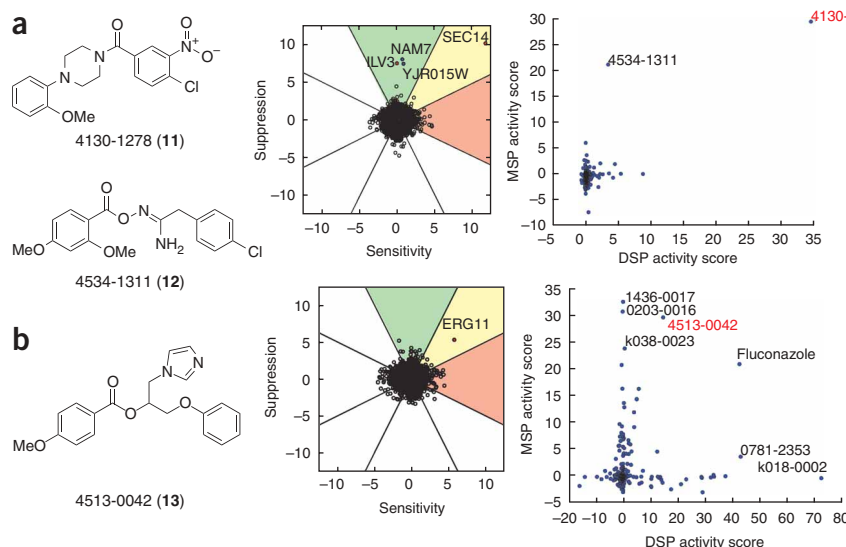

Figure 5 Characterization of new bioactive compounds. (**a**) Left, chemical structure of 4130-1278; middle, MSP versus DSP plot for 4130-1278; right, MSP activity score versus DSP activity score for *SEC14* for all 196 compounds. Datum for 4130-1278 is highlighted in red. (**b**) Left, chemical structure of 4513-0042; middle, MSP versus DSP plot for 4513-0042; right, MSP activity score versus DSP activity score for *ERG11* for all 196 compounds. Data for 4513-0042 are highlighted in red.

We next mapped the ε scores of strains described in **Figure 4c** (sensitive to either compound or the cocktail) to a network comprising 15 types of interactions as defined by BioGRID[38] (**Fig. 4d**). We found that particular neighborhoods of the network were enriched with genes having negative ε scores. In addition, searching this network for highly connected modules of protein-protein interactions (see Methods) identified several modules with an average ε score of less than zero (**Fig. 4e**). These results illustrate that drug cocktails can specifically target genetically unique cells within a heterogeneous population.

Characterization of new bioactive compounds

To demonstrate the utility of these integrated assays in understanding the bioactivity of a more challenging chemical set, we screened 188 synthetic chemical compounds with uncharacterized activities. These compounds were selected from a library of compounds that have
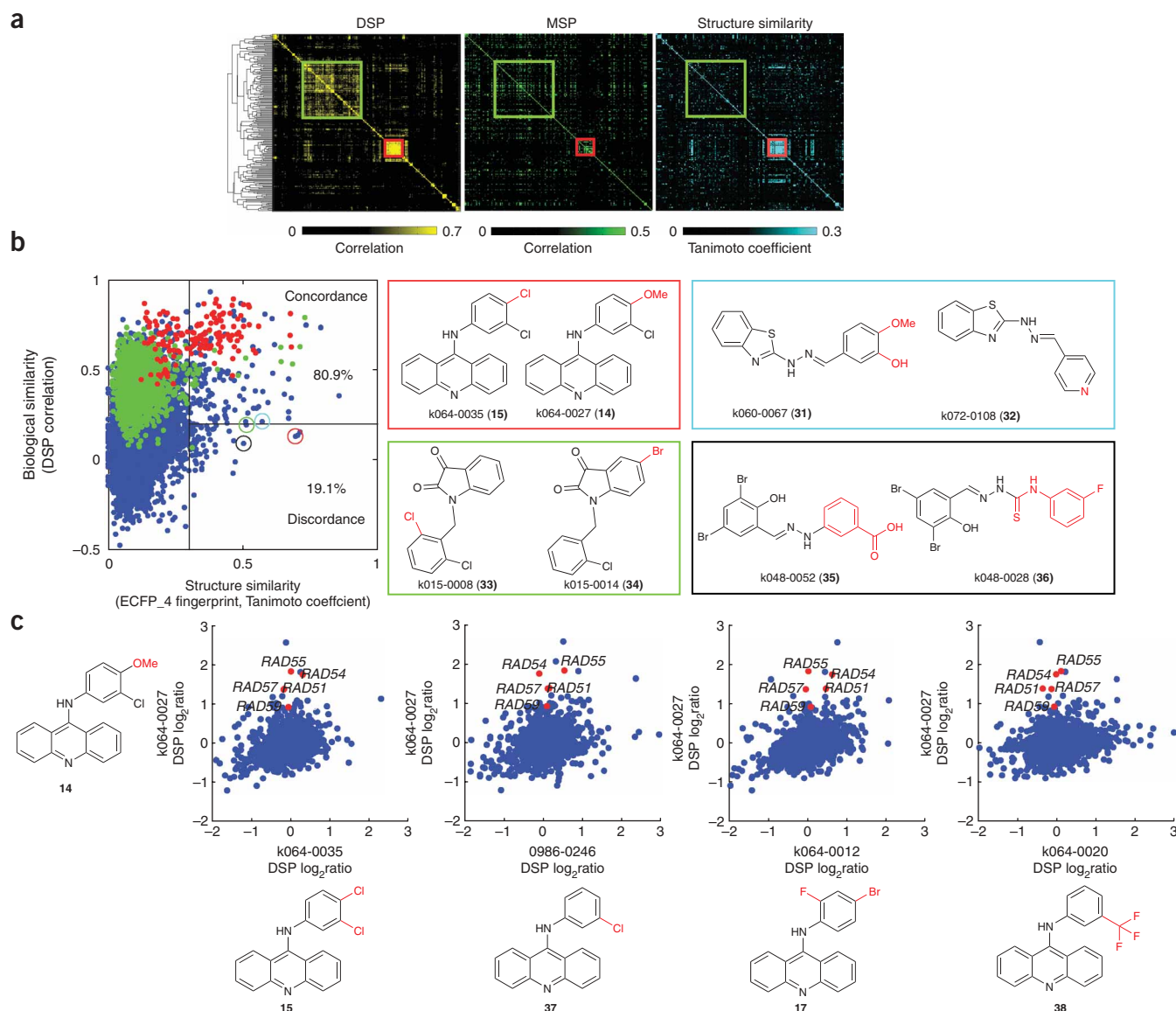
Figure 6 Correlation of chemogenomic profiles with compound structure similarity. (**a**) Similarity matrices, represented as heatmaps, were generated for DSP, MSP and chemical structure similarity for 196 compounds. Chemogenomic profile similarity was determined using Pearson's correlation and colored black to yellow for DSP or black to green for MSP. Structure similarity was determined by Tanimoto similarities in ECFP_4 fingerprints and colored black to blue. Matrices are ordered based on hierarchical clustering of DSP profiles. The dendrogram for the clustering is shown on the left. Two classes of compounds pairs are highlighted in boxes: (i) compounds that share profile correlation but not structural similarity (green) and (ii) compounds that share profile correlation and structural similarity (red). (**b**) Structure-activity concordance was assessed by comparing Tanimoto similarities between each pair of compound to their DSP profile similarity. Compound pairs corresponding to the green and red boxes in **a** are plotted in green and red, respectively. Compound pairs with high structural similarity (Tanimoto score ≥ 0.3) and high profile correlation (correlation > 0.2) are considered concordant, whereas compound pairs with high structural similarity and low profile correlation (correlation ≤ 0.2) are considered discordant. Representative discordant compound pairs (circled red, cyan, green and black) and their corresponding compound structures are shown (matching colored boxes). (**c**) Pairwise comparison of DSP profiles (log₂ ratios) of k064-0027 with four other structurally similar compounds. Genes involved in homologous recombination-mediated DNA repair (*RAD51, RAD54, RAD55, RAD57* and *RAD59*) are highlighted in red.

Table 2 ECFP_4 Tanimoto coefficient similarity of 9-anilinoacridine derivatives to *m*-AMSA

Name	0986-0246	k064-0012	k064-0020	k064-0027	k064-0035	*m*-AMSA
0986-0246	1.00	0.43	0.62	0.57	0.67	0.38
k064-0012		1.00	0.43	0.38	0.43	0.42
k064-0020			1.00	0.47	0.54	0.36
k064-0027				1.00	0.69	0.53
k064-0035					1.00	0.41
m-AMSA						1.00

inhibitory activity against yeast. An activity score (**Supplementary Methods**) designed to identify genes for which copy number changes conferred potent and specific sensitivity in DSP and potent and specific resistance in MSP was calculated for each compound-gene pair, thus yielding a matrix of $>10^6$ activity scores. Many of the known interactions from our reference compounds were among the highest scores in this matrix (**Table 1**), which suggests that the reference compounds were specific for their targets. The highest activity score among the synthetic compounds was for 4130-1278 (**11**) and the gene *SEC14*, which encodes a conserved phosphatidylinositol/phosphatidylcholine transfer protein (**Table 1** and **Fig. 5a**). These screen results were confirmed with isogenic cultures and suggest that Sec14 is a target of 4130-1278 (**Supplementary Fig. 7a** online). Moreover, *SEC14* did not score highly in either MSP or DSP with the other 195 compounds we screened, which indicates that 4130-1278 was unique in its ability to target *SEC14* (**Fig. 5a**). We also note that the second highest scoring compound with activity against *SEC14*, 4534-1311 (**12**), shared moderate structural similarity with 4130-1278 (ECFP_4 Tanimoto coefficient, 0.2).

We also identified a potential new inhibitor of Erg11: 4513-0042 (**13**) (**Table 1** and **Fig. 5b**). 4513-0042 contains an azole ring, a hallmark of many characterized Erg11 inhibitors, and specifically identified *ERG11* in both MSP and DSP assays (**Fig. 5b**). Comparing the DSP and MSP activity score for *ERG11* across all 196 compounds revealed that a number of compounds displayed anti-Erg11 activity in either MSP or DSP alone but not in both. Only 4513-0042 and fluconazole displayed activity against Erg11 in both assays.

Comparison of profiles to compound structure

We previously showed that compounds that share a common substructure have similar chemogenomic profiles[1]. More recently, a large-scale study showed that phenotype-based structure-activity relationships (SAR) can be derived using cytological phenotypes from high-content screening[39]. We therefore sought to determine to what extent chemical structural similarities are reflected in our DSP and MSP results. For each compound, a circular molecular fingerprint was defined using ECFP_4 descriptors. A similarity matrix based on Tanimoto scores was used to describe the pair-wise relationships between the 196 compounds screened. We also generated chemogenomic profile similarity matrices, one for DSP and one for MSP, by calculating the Pearson's correlation between each pair of experiments. Each matrix is presented as a heatmap with compounds ordered by hierarchical clustering of the DSP profiles (**Fig. 6a**). We found the relationship between chemical structure and profiling results to be significant for both DSP ($P < 5.53 \times 10^{-13}$) and MSP ($P < 1.2 \times 10^{-5}$), with DSP exhibiting a higher correlation with structure than MSP (**Supplementary Fig. 7b**). Nevertheless, it is evident from the heatmaps that while some structurally similar chemicals produced similar DSP and/or MSP results (red boxes), other structurally dissimilar compounds also produced highly correlated profiles (green boxes).

To examine how small changes in chemical structure can affect function or bioactivity, we compared Tanimoto similarities with DSP profile similarities between each compound pair. Using a Tanimoto structural similarity score cutoff of ≥ 0.3, we found that ~80% of the compound pairs with similar structures show concordance in their chemogenomic profiles, whereas ~20% of the compound pairs show discordance (**Fig. 6b**). Four exemplar discordant compound pairs are shown in **Figure 6b**. Compounds k064-0027 (**14**) and k064-0035 (**15**), for example, differ only in one position, but they differ greatly in their respective DSP profiles. These compounds are structurally similar to a family of 9-anilinoacridine derivatives that includes amsacrine (*m*-AMSA, **16**) (and its analogs); the derivatives have been extensively studied as potential antitumor agents[40]. *m*-AMSA has been shown to interact with DNA and also to inhibit topoisomerase II (ref. 41). k064-0027 shared the greatest structural similarity with *m*-AMSA among these derivatives (**Table 2**), and comparison of DSP results for k064-0027, k064-0035 and three other derivatives in our dataset showed that strains deficient in DNA repair are sensitive only to k064-0027 (**Fig. 6c**). This observation is consistent with the methyl ester group unique to k064-0027, as quantitative structure-activity relationship (QSAR) analysis of 643 members of the 9-anilinoacridine family showed that electron-donating groups are favored on the aniline ring[42]. Two of the compounds that do not induce a DNA damage response (k064-0012 (**17**) and k064-0035) both contain two halogen substituents, and it is known from the QSAR study[42] that electron-withdrawing groups reduce antitumor activity for this family of compounds. In addition to having an electron-donating group instead of an electron-withdrawing group, k064-0027 is also less lipophilic, having a calculated logP of 5.5 compared with 6.14 for k064-0035. For comparison, *m*-AMSA has a calculated logP of 4.00.

DISCUSSION

Bioactive chemical probes are useful biological tools and can be as effective as mutants or antibodies for studying the functions of genes and pathways relevant in human health. Here we describe an integrated, miniaturized platform that comprehensively interrogates the effect of increasing and decreasing individual gene dosage on drug resistance using a single microarray to measure cellular fitness. Our experiments with eight reference compounds and 188 compounds of unknown activity validated the use of this platform for identifying small molecule–target interactions. From a practical perspective, extensibility is an essential feature of genome-wide assays, and accordingly, we designed this suite of assays to be compatible with existing protocols, reagents and devices. Additionally, miniaturization is important for high-throughput screens where compounds, in particular natural products, are limited in supply.

Much of the current cost of drug development can be attributed to a poor understanding of the effects of potential drugs *in vivo*. The practice of target-based drug discovery has indeed produced many promising drug candidates that ultimately fail clinical trials due to unanticipated side effects[43]. Systems-level approaches such as those described here that measure small-molecule effects in an *in vivo* context offer to better define the biological effects of potential drugs. For example, the protein phosphatase inhibitors cantharidin and calyculin A have similar activities *in vitro*; however, their deletion

sensitivity and multicopy suppression profiles are surprisingly dissimilar (**Fig. 3**). An additional benefit of our approach is that all potential targets are interrogated simultaneously and without bias. This approach readily identified the therapeutic targets of several drugs (**Fig. 2**). Our efforts with 188 previously uncharacterized synthetic compounds, which we expect will be a valuable resource for future experiments (and analysis), identified several potential interactions with various cellular targets (**Fig. 5**). Though further chemical modifications would be needed to confer the pharmacokinetic properties required for a therapeutic, compounds identified using this methodology could prove to be excellent starting points for developing analogs with improved properties. In principle, these modifications could be directed in part by SAR studies similar to those illustrated in **Figure 6**.

Using a randomly generated genomic library for MSP has several key benefits. One is that it is readily adaptable to testing the genomes of other organisms, provided an appropriate expression array is available. We demonstrated this by identifying a putative functional homolog of an uncharacterized and poorly conserved *S. cerevisiae* gene (*CRG1*) in the fungal pathogen *C. albicans* (**Fig. 3e,f**). Given the conservation between yeast and human genes, expression of human complementary DNAs in yeast could prove useful in identifying human drug targets.

We also demonstrate the utility of chemogenomic profiles for the study of drug combinations. Understanding how drug combinations affect cell physiology is essential for three reasons: (i) to predict adverse drug interactions, (ii) to develop better strategies for designing effective drug combination treatment regimens and (iii) to develop chemical genetic strategies for studying cellular processes. Recent advances in 'combination high-throughput screening' (cHTS)[44,45] promise the identification of many drug combinations that result in non-additive interactions. However, determining the links between the observed drug interaction and the underlying molecular mechanism is not straightforward. As a step toward this goal, we showed that a genome-wide analysis of deletion strains with drug combinations in the context of the underlying genetic network enables the identification of processes that buffer the effects of drug-drug combinations. Although not investigated in this study, we also foresee the use of MSP to identify suppressors of drug interactions.

In summary, integration of multiple genome-wide assays is an effective way to use small molecules to dissect fundamental cellular processes. Development of scalable technologies in addition to those described here will be important—as will the development of analysis tools that can integrate these genome-wide datasets to piece together the complete spectrum of chemical genetic interactions.

METHODS

Reagents. Methotrexate, rapamycin, MMS, nocodazole, cantharidin and fluconazole were purchased from Sigma-Aldrich. Calyculin A and latrunculin A were purchased from BioMol. 1-NM-PP1 was chemically synthesized as previously described[46]. The chemical diversity library was obtained from ChemDiv. Each compound was dissolved in DMSO, with the exception of rapamycin, which was dissolved in 90% ethanol, 10% Tween 20. All compounds were stored at –20 °C until use.

Deletion pool construction and screening conditions. Deletion pool construction was carried out as previously described[3,12] with the following modification. Only strains identified as essential for growth in rich medium, sterile or deficient in mating were used to create the heterozygous deletion pool. Pooled growth of the homozygous deletion pool and the heterozygous deletion pool was carried out for 5 and 20 generations, respectively. Both pools were inoculated at an optical density at 600 nm (OD_{600}) of 0.02 and grown in 48-well microtiter plates (Nunc) in a volume of 700 μl and in compound concentrations that inhibited pool growth by ∼10%. Cells were harvested by a Packard Multiprobe II four-probe liquid-handling system (PerkinElmer). For 20-generation experiments, cells were maintained in logarithmic phase by robotically diluting cultures every five doublings.

Multicopy pool construction and screening conditions. An *S. cerevisiae* random genomic library constructed in a high-copy 2 μM expression vector (YEplac195) was transformed into yeast (*cdc28-as* or BY4743) by a standard lithium acetate method[47] and selected on medium lacking uracil (URA⁻). After 3 d of growth, ∼10^6 transformants were pooled into medium containing 7% DMSO, aliquoted and stored at –80 °C until use. A *C. albicans* genomic library constructed in a high-copy 2 μM expression vector (pRS426)[48] was purchased from Open Biosystems. For both pools, frozen aliquots were thawed and inoculated directly into URA⁻ medium to an OD_{600} of 0.02 and a volume of 700 μl. Compound was added, and the pool was grown for five generations in 48-well microtiter plates (Nunc) at inhibitory concentrations of at least 50% (IC_{50}) and harvested the same way as described above.

Microarray analysis. Both DSP and MSP were analyzed using a high-density oligonucleotide tag array manufactured by Affymetrix[13]. MSP for *C. albicans* was analyzed using a custom high-density oligonucleotide genechip also manufactured by Affymetrix[49] (PN = 510556). For both DSP and MSP, barcode probe intensities were extracted and processed as previously described[12]. Each array was mean normalized, and fold change (log_2 control/treatment) was calculated by comparing to a set of control arrays. Tags from the homozygous pool were normalized separately from tags from the heterozygous essential pool, as were the upstream tags (uptag) and downstream tags (downtag). At least two biological replicates were performed for each of the eight reference conditions. The log_2 ratios of both tags were averaged to generate a single score for each gene. For MSP, log_2 ratios of intensities were ordered by each ORF's genomic location and analyzed using a sliding window to identify loci that had at least two adjacent ORFs with log_2 ratios ≥ 1.6. Statistical analyses for identifying significant strains and activity scores are further described in the **Supplementary Methods**. All raw and analyzed data are available on the author's website (http://chemogenomics.stanford.edu/HIP/supplements/05chemo/).

Quantifying dose-response matrix drug interactions. For each drug, 1.25-fold serial drug dilutions were performed in DMSO. Wild-type yeast (BY4743) were grown overnight to saturation and diluted into YPD (yeast extract, peptone, dextrose) medium to an OD_{600} of 0.2 and aliquoted in an 8×8 matrix (96-well plate). Each drug combination pair was then added to yield the dose-response matrix and grown using Tecan GENios microplate readers for up to 30 h. At least two replicates were conducted for each growth condition. We used area under the growth curve (AUGC) as a metric to capture both defects in growth rate and carrying capacity. Each dose matrix was scored for drug interaction $\varepsilon = AUGC_{i,j} - AUGC_i AUGC_j$, where $AUGC_{i,j}$ represents the growth defect conferred by treatment with drug i and drug j, and $AUGC_i$ and $AUGC_j$ represent the growth defect conferred by single treatment of $drug_i$ and $drug_j$, respectively, at the same concentration normalized to no drug growth. The ε values for each replicate were averaged and the dose matrix was represented as a heatmap using Matlab (MathWorks). For analysis of the morphology of the response surfaces, we used shape models as previously described[37]. Detailed methods describing this analysis can be found in the **Supplementary Methods**.

Quantifying deletion strain drug interactions. We used rankproduct analysis[50] to identify strains that were sensitive to each compound and the cocktail (**Fig. 4b**). Log_2 ratios were calculated pairwise for each control, and treatment array and rank were sorted. The product of the ranks for each pair of arrays was calculated. Significance was estimated using a permutation-based procedure. Strains meeting a false discovery rate of 0.1 were mapped to the interaction network shown in **Figure 4d**. For each strain x, the Z-scores for each condition were then used to calculate $\varepsilon_x = W_{x,cantharidin, calyculin \ A} - W_{x,cantharidin} \cdot W_{x,calyculin \ A}$ to characterize the degree of drug interaction in that strain, where $W_{x,y}$ represents the Z-score for the log_2 ratio for strain x in condition y. This

was performed separately for the heterozygous and homozygous pools to normalize for differences in the ratios between the two pools. We filtered strains for which ε could not be accurately assigned because the strain was highly sensitive to either single drug condition or had low overall TAG intensity in the control. To identify these strains, we calculated for each strain the maximum possible ε value by calculating $\varepsilon_{max,x} = Bgd - W_{x,cantharidin} \cdot W_{x,calyculin\ A}$, where Bgd is the background intensity ($\sim$60) simulating that the strain is highly sensitive to the cocktail. Strains for which ε_{max} is negative were excluded from the analysis and are represented as gray nodes in **Figure 4d**.

Chemical descriptor annotation. ECFP_4 fingerprints were calculated using Pipeline Pilot V6.1.1 (Scitegic), and logP values were calculated using Molecular Operating Environment version 2007.09 (Chemical Computing Group, Inc).

Note: Supplementary information and chemical compound information is available on the Nature Chemical Biology website.

AUTHOR CONTRIBUTIONS
C.N., R.P.S. and S.H. designed the study, analyzed the data and wrote the paper; S.H. and R.P.S. did the experiments. G.G. analyzed the data and helped write the paper. A.S. did the *glc7* allele analysis; I.M.W. did the cheminformatic analysis. M.P. designed the robotic assay platform; E.F. designed the database infrastructure. K.M.S. and C.Z. designed the *cdc28-as* experiment. R.W.D. provided valuable advice. M.M. and S.S. helped with genome-wide screens.

Published online at http://www.nature.com/naturechemicalbiology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Giaever, G. *et al.* Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc. Natl. Acad. Sci. USA* **101**, 793–798 (2004).
2. Lum, P.Y. *et al.* Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* **116**, 121–137 (2004).
3. Lee, W. *et al.* Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS Genet.* **1**, 235–246 (2005).
4. Parsons, A.B. *et al.* Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* **126**, 611–625 (2006).
5. Rine, J., Hansen, W., Hardeman, E. & Davis, R.W. Targeted selection of recombinant clones through gene dosage effects. *Proc. Natl. Acad. Sci. USA* **80**, 6750–6754 (1983).
6. Luesch, H. *et al.* A genome-wide overexpression screen in yeast for small-molecule target identification. *Chem. Biol.* **12**, 55–63 (2005).
7. Butcher, R.A. *et al.* Microarray-based method for monitoring yeast overexpression strains reveals small-molecule targets in TOR pathway. *Nat. Chem. Biol.* **2**, 103–109 (2006).
8. Parsons, A.B. *et al.* Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat. Biotechnol.* **22**, 62–69 (2004).
9. Hughes, T.R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126 (2000).
10. Kung, C. *et al.* Chemical genomic profiling to identify intracellular targets of a multiplex kinase inhibitor. *Proc. Natl. Acad. Sci. USA* **102**, 3587–3592 (2005).
11. Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
12. Pierce, S.E. *et al.* A unique and universal molecular barcode array. *Nat. Methods* **3**, 601–603 (2006).
13. Pierce, S.E., Davis, R.W., Nislow, C. & Giaever, G. Genome-wide analysis of barcoded *Saccharomyces cerevisiae* gene-deletion mutants in pooled cultures. *Nat. Protoc.* **2**, 2958–2974 (2007).
14. Brenner, C. Chemical genomics in yeast. *Genome Biol.* **5**, 240 (2004).
15. Giaever, G. *et al.* Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat. Genet.* **21**, 278–283 (1999).
16. Li, X. *et al.* Multicopy suppressors for novel antibacterial compounds reveal targets and drug efflux susceptibility. *Chem. Biol.* **11**, 1423–1430 (2004).
17. Breitling, R., Armengaud, P. & Amtmann, A. Vector analysis as a fast and easy method to compare gene expression responses between different experimental backgrounds. *BMC Bioinformatics* **6**, 181 (2005).
18. Myers, C.E., Lippman, M.E., Elliot, H.M. & Chabner, B.A. Competitive protein binding assay for methotrexate. *Proc. Natl. Acad. Sci. USA* **72**, 3683–3686 (1975).
19. Kontoyiannis, D.P., Sagar, N. & Hirschi, K.D. Overexpression of Erg11p by the regulatable GAL1 promoter confers fluconazole resistance in *Saccharomyces cerevisiae. Antimicrob. Agents Chemother.* **43**, 2798–2800 (1999).
20. Zheng, X.F., Florentino, D., Chen, J., Crabtree, G.R. & Schreiber, S.L. TOR kinase domains are required for two distinct functions, only one of which is inhibited by rapamycin. *Cell* **82**, 121–130 (1995).
21. Koltin, Y. *et al.* Rapamycin sensitivity in *Saccharomyces cerevisiae* is mediated by a peptidyl-prolyl cis-trans isomerase related to human FK506-binding protein. *Mol. Cell. Biol.* **11**, 1718–1723 (1991).
22. Honkanen, R.E. Cantharidin, another natural toxin that inhibits the activity of serine/threonine protein phosphatases types 1 and 2A. *FEBS Lett.* **330**, 283–286 (1993).
23. Li, Y.M. & Casida, J.E. Cantharidin-binding protein: identification as protein phosphatase 2A. *Proc. Natl. Acad. Sci. USA* **89**, 11867–11870 (1992).
24. Ishihara, H. *et al.* Calyculin A and okadaic acid: inhibitors of protein phosphatase activity. *Biochem. Biophys. Res. Commun.* **159**, 871–877 (1989).
25. Fujiki, H. & Suganuma, M. Tumor promotion by inhibitors of protein phosphatases 1 and 2A: the okadaic acid class of compounds. *Adv. Cancer Res.* **61**, 143–194 (1993).
26. MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P. & Codd, G.A. Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1 and 2A from both mammals and higher plants. *FEBS Lett.* **264**, 187–192 (1990).
27. Walsh, E.P., Lamont, D.J., Beattie, K.A. & Stark, M.J. Novel interactions of *Saccharomyces cerevisiae* type 1 protein phosphatase identified by single-step affinity purification and mass spectrometry. *Biochemistry* **41**, 2409–2420 (2002).
28. Krogan, N.J. *et al.* High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell* **13**, 225–239 (2004).
29. Andrews, P.D. & Stark, M.J. Type 1 protein phosphatase is required for maintenance of cell wall integrity, morphogenesis and cell cycle progression in *Saccharomyces cerevisiae. J. Cell Sci.* **113**, 507–520 (2000).
30. Baker, S.H., Frederick, D.L., Bloecher, A. & Tatchell, K. Alanine-scanning mutagenesis of protein phosphatase type 1 in the yeast *Saccharomyces cerevisiae. Genetics* **145**, 615–626 (1997).
31. Hsu, J.Y. *et al.* Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* **102**, 279–291 (2000).
32. Bloecher, A. & Tatchell, K. Defects in *Saccharomyces cerevisiae* protein phosphatase type I activate the spindle/kinetochore checkpoint. *Genes Dev.* **13**, 517–522 (1999).
33. Hillenmeyer, M.E. *et al.* The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362–365 (2008).
34. Martin, J.L. & McMillan, F.M. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol.* **12**, 783–793 (2002).
35. Arnaud, M.B. *et al.* Sequence resources at the Candida Genome Database. *Nucleic Acids Res.* **35**, D452–D456 (2007).
36. Bliss, C. The toxicity of poisons applied jointly. *Ann. Appl. Biol.* **26**, 585–615 (1939).
37. Lehar, J. *et al.* Chemical combination effects predict connectivity in biological systems. *Mol. Syst. Biol.* **3**, 80 (2007).
38. Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).
39. Young, D.W. *et al.* Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nat. Chem. Biol.* **4**, 59–68 (2008).
40. Baguley, B.C. & Nash, R. Antitumour activity of substituted 9-anilinoacridines–comparison of in vivo and in vitro testing systems. *Eur. J. Cancer* **17**, 671–679 (1981).
41. Nelson, E.M., Tewey, K.M. & Liu, L.F. Mechanism of antitumor drug action: poisoning of mammalian DNA topoisomerase II on DNA by 4′-(9-acridinylamino)-methanesulfon-m-anisidide. *Proc. Natl. Acad. Sci. USA* **81**, 1361–1365 (1984).
42. Denny, W.A. *et al.* Potential antitumor agents. 36. Quantitative relationships between experimental antitumor activity, toxicity, and structure for the general class of 9-anilinoacridine antitumor agents. *J. Med. Chem.* **25**, 276–315 (1982).
43. Terstappen, G.C., Schlupen, C., Raggiaschi, R. & Gaviraghi, G. Target deconvolution strategies in drug discovery. *Nat. Rev. Drug Discov.* **6**, 891–903 (2007).
44. Zimmermann, G.R., Lehar, J. & Keith, C.T. Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov. Today* **12**, 34–42 (2007).
45. Borisy, A.A. *et al.* Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. USA* **100**, 7977–7982 (2003).
46. Bishop, A.C. *et al.* A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000).
47. Gietz, R.D., Schiestl, R.H., Willems, A.R. & Woods, R.A. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355–360 (1995).
48. Kadosh, D. & Johnson, A.D. Rfg1, a protein related to the *Saccharomyces cerevisiae* hypoxic regulator Rox1, controls filamentous growth and virulence in *Candida albicans. Mol. Cell. Biol.* **21**, 2496–2505 (2001).
49. Lan, C.Y. *et al.* Metabolic specialization associated with phenotypic switching in *Candida albicans. Proc. Natl. Acad. Sci. USA* **99**, 14907–14912 (2002).
50. Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* **573**, 83–92 (2004).

CORRIGENDA

CORRIGENDUM: An integrated platform of genomic assays reveals small-molecule bioactivities

Shawn Hoon, Andrew M Smith, Iain M Wallace, Sundari Suresh, Molly Miranda, Eula Fung, Michael Proctor, Kevan M Shokat, Chao Zhang, Ronald W Davis, Guri Giaever, Robert P St Onge & Corey Nislow

In the version of this article initially published, there was a space missing in the author name Robert P St Onge. The error has been corrected in the HTML and PDF versions of the article.