www.sciencemag.org/344/6180/208/suppl/DC1

Supplementary Materials for

Mapping the Cellular Response to Small Molecules Using Chemogenomic Fitness Signatures

Anna Y. Lee, Robert P. St.Onge, Michael J. Proctor, Iain M. Wallace, Aaron H. Nile, Paul A. Spagnuolo, Yulia Jitkova, Marcela Gronda, Yan Wu, Moshe K. Kim, Kahlin Cheung-Ong, Nikko P. Torres, Eric D. Spear, Mitchell K. L. Han, Ulrich Schlecht, Sundari Suresh, Geoffrey Duby, Lawrence E. Heisler, Anuradha Surendra, Eula Fung, Malene L. Urbanus, Marinella Gebbia, Elena Lissina, Molly Miranda, Jennifer H. Chiang, Ana Maria Aparicio, Mahel Zeghouf, Ronald W. Davis, Jacqueline Cherfils, Marc Boutry, Chris A. Kaiser, Carolyn L. Cummins, William S. Trimble, Grant W. Brown, Aaron D. Schimmer, Vytas A. Bankaitis, Corey Nislow, Gary D. Bader, Guri Giaever*

*Corresponding author. E-mail: g.giaever@ubc.ca

Published 11 April 2014, *Science* **344**, 208 (2014)
DOI: 10.1126/science.1250217

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/344/6180/208/suppl/DC1)

Materials and Methods

All computational analysis was performed using R (*24*), unless otherwise indicated.

Yeast chemical screening

We designed a chemical screening protocol and, with few exceptions, used this protocol for each of the 3,356 genome-wide screens described in this study. Frozen aliquots of heterozygous pool (~1,095 strains) and homozygous pool (~4,810 unique strains) were thawed and separately diluted to an OD_{600} of 0.03125 in 700µL of screening media (YPD + 25mM HEPES, pH 6.8). Diluted pools were grown in 48-well plates for 4hrs at 30ºC in a Tecan Genios (Tecan Systems, Inc. San Jose, CA) spectrophotometer, at which point 7µL of a 100X concentrated stock of the screening compound (in DMSO, the diluent for the chemicals in this study) was added to the culture using a Packard Multiprobe (Perkin Elmer, Santa Clara, CA) liquid handler. Homozygous pool cultures were grown in duplicate wells on the same plate for 5 doublings, automatically collected and stored at 4ºC in a 48-well plate for no longer than 2 days. Heterozygous cultures were grown for 20 doublings and were maintained in logarithmic growth by automated dilution of the cultures every 5 generations into wells containing only media and compound. After 20 doublings, the heterozygous samples were added to the collected homozygous pool samples and stored for no longer than 1 day at 4ºC. Combined pool samples were resuspended by manual pipetting and transferred from the 48-well plate to a microfuge tube, and pelleted at 14,000rpm for 3min. Supernatant was removed and the pellets stored at -20ºC. Cell pellets were thawed, resuspended in 100µL of cell digestion buffer (Zymo Research, Irvine, CA) and 5µLs of R-zymo (Zymo Research, Irvine, CA) and incubated at 37ºC for 1h. Genomic DNA was extracted from each sample using a Qiaxtractor (Qiagen). PCR amplification of molecular tags in 96-well format, followed by microarray hybridization, washing, and scanning were performed as previously described (*25, 26*).

We define an 'experimental group' as a set of 96 or fewer samples whose tags were PCR amplified concurrently in the same plate. By design, each experimental group consists of 84 chemical-treatment samples and 12 control samples (in which pools were treated with 1% DMSO). Efforts were made to minimize technical variability within an experimental group. In addition to concurrent PCR amplification of tags, experiments within a group were grown in media prepared on the same date and samples from the same group were hybridized to tag arrays in succession (typically 14 arrays per day over the course of ~10 days). Fitness defect (FD) scores for each deletion strain were derived from the ratio of a tag's signal in the treatment samples and its signal in the control samples (see below for details).

Mammalian viability assays

For the TEX human leukemia cells (a gift from Dr. J. Dick, Toronto, Canada), AML2 acute myeloid leukemia line and LP1 myeloma cell line, growth and viability were measured with the CellTiter-Fluor assay (Promega, Madison, WI, USA) as per manufacturer's instructions and by trypan blue staining.

The HCT116 colorectal cell line was propagated in McCoy's media with 10% FBS, seeded at 2500 cells/well. 6h after seeding, compounds were added by diluting each compound in McCoy's media to a final concentration of 1mM and 1% DMSO was then added to cells such that the DMSO concentration was fixed at 0.5%. 72h post treatment, cells were fixed by addition

of formaldehyde directly to the culture media at a final concentration of 1%. Fixed cells were washed twice in PBS, and stained by addition of Hoechst 33342 dye (1µg/mL). Cells were imaged using a 4X objective and phase contrast, and nuclear fluorescence images were overlaid. At least 1000 cells of each well were counted manually in a blind manner.

A549 cells (human lung adenocarinoma) were obtained from ATCC (http://www.atcc.org) and were seeded in 96-well plates at 2200 cells/well in 200µL of medium. Compounds were serially diluted first in DMSO then in RPMI, with the final DMSO concentration not exceeding 1%. After 72h of incubation, cell viability was measured with a sulforhodamine B (SRB) viability assay (*27*) as described (*28*).

Raw data normalization and removal of problematic tags

Each probe on the Genflex tag16k array (Affymetrix, Santa Clara, CA), i.e. the Tag4 array, is represented by 5 replicate features. These replicates allow the removal of outliers that may for example, arise from small debris in the hybridization solution. To identify and remove probes defined as outliers, we used a previously described masking algorithm (*29*). We next defined the 'raw average' of each tag as the average of all remaining probe replicates for a particular tag. We then removed all tags corresponding to the control strain.

For each array, uptags and downtags were normalized separately, as were heterozygous and homozygous strains, creating 4 sets: uptag/het, uptag/hom, downtag/het, downtag/hom. For each set, logged raw average intensities were normalized across all arrays using a variation of median polish that incorporates batch effect correction. Briefly each iteration involves:

i) **Array median:** shifting intensities so that the median of each array matches the overall median
ii) **Batch median:** for each tag, shifting intensities so that the median of the tag, across a group batch, matches the median across all arrays
iii) **Batch MAD:** for each tag, scaling intensities so that the median absolute deviation (MAD) of the tag, across a group batch, matches the MAD across all arrays
iv) **Date median:** for each tag, shifting intensities so that median of the tag, across a hybridization date batch, matches the median across all arrays
v) **Date MAD:** for each tag, scaling intensities so that the MAD of the tag, across a date batch, matches the MAD across all arrays
vi) **Pool inhibition:** for each tag, shifting intensities so that the trend line for the level of pool inhibition of each array sample (estimated from the doubling time of the sample), versus intensity, is a horizontal line

We stop the iterative algorithm when the adjustment amount per iteration stops decreasing.

For each array, we computed compound and control background thresholds (compound = <median signal>, control = <median signal> + 5 x <MAD of signal>) using the unnormalized intensity values of the unused features on the array, and then defined overall compound and control thresholds with the medians of these thresholds (across all arrays). A tag is excluded from further analysis if its median normalized intensity in the control arrays is less than the overall control threshold. Such low signal suggests that the tag is faulty, or that the corresponding strain has low abundance even when grown in the absence of a compound. We thus eliminated the tag to avoid inaccurate estimates of the fitness defects of strains.

As each strain was constructed with two unique molecular tags (an 'uptag' and a 'downtag'), and because the performance of these two tags can vary significantly, we identified the 'best tag' of each strain. If only one tag is still available after the control-based background filter, it is the best tag by default. If both tags are available, we defined the best tag as the tag with lowest robust coefficient of variation (<MAD intensity> / <median intensity>) in the control arrays. We filtered out all non-best tags and for each remaining tag, we scaled the intensities so that the MAD across the control arrays matches the median of all such MADs. This step forces all tags to have the same variability and therefore compensates for variability due to properties of the tag (e.g. hybridization variability as opposed to biological variability). Furthermore, all intensity values below the overall compound background threshold were re-assigned to the threshold value. Assuming that intensity measurements are less accurate below the threshold, this step avoids inaccurate estimates of the fitness defects of strains.

To simplify our dataset, we removed strains where the deletions no longer correspond to valid genes according to the Saccharomyces Genome Database (SGD), and strains where the zygosity contradicts the viability of the strain as currently reported in SGD (e.g. a homozygous strain in our pool reported to be inviable). These contradictions potentially arise due to secondary site mutations in the strains in the deletion collection. Furthermore, some genes have multiple deletion strains in our pools and we selected one strain for each gene. In some cases, a strain is deleted for a gene that has been merged with another gene, which already has its own deletion strain. In these cases, we selected the latter strain because the deleted region better matches the currently defined gene region. For each of the remaining genes with multiple strains, we selected the strain that better recapitulates known functional similarities between the deleted gene and all other genes. Specifically, we used the intensities to computed fitness defect (FD) scores (described below), and then computed the cofitness (described below) between each strain of interest and all other strains (resulting in ~6,000 values per strain). For the gene deleted from a strain of interest, we obtained measures of the semantic similarity between it and all other genes (resulting in ~6,000 values per gene), where the values are derived from biological process annotations from Gene Ontology (GO) and therefore reflect known functional similarities (*30*). Since cofitness is a measure of functional similarity based on our chemogenomic dataset, for each strain/gene of interest, we computed the correlation between the cofitness values and the semantic similarity values. For strains deleted for the same gene, we selected the strain with greater cofitness-semantic similarity correlation to represent the gene. The final list of strains/genes included in this study is given in **table S9**.

After computing FD scores for the final strains, we filtered out data points that demonstrate lingering batch effects. Some date batches contained a small number of arrays and it was thus difficult to accurately estimate and correct the corresponding batch effects during normalization. We therefore estimated lingering batch effects by manually examining date batches with <14 arrays (and ≥3 arrays since we cannot properly estimate with fewer). For each strain, we computed the median and MAD FD score across the arrays in each small batch. If the median is significantly positive/negative (standard normal distribution $P < 0.01$) and the MAD < 2, the strain consistently demonstrated altered fitness within the batch. Since we expect fitness defect scores to be centered at zero, we filtered out the data points corresponding to these aberrant strain-batch combinations (i.e. ~0.04% of the dataset).

In determining the best normalization approach, we tried existing approaches such as Supervised Normalization of Microarrays [SNM; (*31*)] and quantile normalization. While we found SNM to be promising due to its flexibility and ability to reduce batch effects in our dataset (**fig. S19**), the runtime slowed the process of optimizing the parameters for our dataset. Comparatively, quantile normalization is fast, yet it does not include explicit batch effect correction. We therefore incorporated batch effect correction by computing reference intensity distributions for each experimental group (from group averages), and quantile normalized profiles in each group to the reference distribution of the group. Moreover, 'best tag' selection was performed separately for each group, and FD scores were computed using group-matched controls, followed by an additional step where, within each group, all scores of a given strain are converted into robust *Z* scores. While this quantile normalization approach reduced batch effects, it did not reduce chip date batch effects as effectively as the median polish variant we ultimately chose to use (**fig. S19**).

The fitness defect (FD) score

We devised a fitness defect (FD) score that quantifies the sensitivity of each deletion strain to a chemical perturbation by comparing the signal of a strain following chemical treatment to the signal of the strain from control samples (i.e. DMSO-treated samples). Specifically, $\log_2$ratios were calculated for each strain as follows:

$\log_2\text{ratio} = \log_2[\langle\text{median signal from control samples}\rangle / \langle\text{signal from chemical sample}\rangle]$

To facilitate comparisons between chemical treatments, i.e. chemogenomic screens, we standardized the $\log_2$ratios (separately for heterozygous and homozygous strains). Specifically, the FD score of strain *i* in screen *j* was computed as follows:

$FD_{i,j} = (\log_2\text{ratio}_{i,j} - \langle\text{median of } \log_2\text{ratios for screen } j\rangle) / \langle\text{MAD of } \log_2\text{ratios for screen } j\rangle$

FD scores are essentially robust *Z* scores, and for each screen, the median FD score is 0 and the MAD is 1. The scores are used throughout the manuscript and are searchable on our webpage (http://chemogenomics.pharmacy.ubc.ca/ycs/).

For a given screen, combining the FD scores of heterozygous strains (from HIP) with the scores of homozygous strains (from HOP) results in a HIPHOP profile of the chemical treatment.

Identification of significant chemical-genetic interactions

We defined significant chemical-genetic interactions by identifying FD scores that deviated significantly from other FD scores in a given screen (heterozygous and homozygous strains were considered separately). This approach is based on the assumption that, at chemical concentrations that only minimally inhibit growth of the pool, most strains will not exhibit a fitness defect. As the FD scores in a screen generally follow a standard normal distribution, we obtained the probability that a given score is an outlier in this distribution (one-sided *P*). If a score has $P < 0.001$, we identify a significant chemical-genetic interaction between the screened chemical and the deleted gene associated with the score. This was the method used to define significant chemical-genetic interactions throughout this study; however, a strain-centric approach is also presented on the webpage. The *P* described above identifies outlier genes for a given chemical, whereas a strain-centric *P* identifies outlier chemicals for a given strain/gene.

Thus, a strain-centric *P* allows users equipped with a gene of interest to identify chemicals exhibiting the strongest interactions with the gene, relative to all other chemicals. Briefly, all FD scores of a given strain are converted into *Z* scores to obtain strain-centric *P* values from the standard normal distribution.

Identification of HIP hits

HIP exploits drug-induced haploinsufficiency, a phenotype whereby a growth inhibitory compound or drug induces a specific fitness defect in a heterozygous strain deleted for one copy of the drug's target (*8*). We therefore identified 'HIP hits' as potential targets of profiled compounds. Because we are most interested in proteins that are targeted with high specificity, we first developed a measure of specificity called 'clearance', which identifies hits that are substantially more significant than all other hits. Furthermore, we are only interested in strains exhibiting fitness defects, so we only computed clearance for strains with FD score > 0 in a given HIP profile. Specifically, the clearance of the strain with the *i*th greatest FD score is defined as follows:

clearance' = $FD_{(i)} - FD_{(i+1)}$ [Note: computed for all strains with FD > 0 before the next step.]

and if $FD_{(i)} \geq FD_{max}$, clearance = $clearance_{max}$

otherwise, clearance = clearance'

where

$FD_{(i)}$ is the *i*th greatest FD score in the profile,

FD_{max} is the FD score of the strain with clearance' = $clearance_{max}$, and

$clearance_{max}$ is the maximum clearance' associated with the profile.

The strain with clearance' = $clearance_{max}$, $strain_{maxC}$, divides the heterozygous strains into two sets where one set (including $strain_{maxC}$) exhibits FD scores greater than the second set by at least $clearance_{max}$ units. For example, $strain_{maxC}$ = *ERG11* in one fluconazole HIP profile (**Fig. 1A**), and it is the only member of the high-scoring set, which exhibits greater fitness defects than the lower-scoring set by ~15.38 units (out of a maximum of 22.16 units in this profile). If $strain_{maxC}$ exhibits a significant FD score (standard normal distribution $P < 0.001$) and $clearance_{max} \geq 5.75$, all genes associated with the high-scoring set are designated HIP hits.

To optimize our threshold, we considered $clearance_{max}$ thresholds from 2.25 to 15 (steps of 0.25), and with each threshold, we identified compound-HIP hit interactions and estimated the significance with which these interactions are enriched with the benchmark compound-target interactions established in yeast (**table S2**). Specifically, 31 established interactions formed our positive gold standard set (involving 26 compounds), 500 randomly selected genes for each compound formed our negative gold standard set, and for each threshold, we counted the number of gold standard compounds where at least one of the associated genes (i.e. known targets for positives, random genes for negatives) was identified as a HIP hit. It should be noted that an established drug-target can be missed by HIP due to suboptimal inhibitory concentrations (e.g. SGTC_227/fluconazole) and/or concentration-dependence of the interaction (*5*). The hypergeometric test was then used to compute a *P* value for the significance with which compounds with HIP hits are enriched for compounds where the HIP hits include a known target. We repeated this computation 100 times with different random samples (re-samples) of the negative gold standards, and re-samples of representative profiles for positive gold standard compounds that were profiled more than once. We selected 5.75 as the clearance threshold

because it yielded the most significant enrichment, measured by the median *P* across the samples. We refer to the compounds in the resulting compound-HIP hit interactions as 'chemical-genetic probes'.

Although the HIP assay, and by extension the HIP hit list, is restricted to essential genes, we note in the HOP profile of fluvastatin (SGTC_2753) that *HMG1*, one of the established non-essential targets of this drug, is associated with highly significant and specific fitness defects. This result demonstrates that HOP profiles can also guide the identification of targets.

Assessing the promiscuity of chemical-genetic probes

To roughly assess the promiscuity of chemical-genetic probes, we extracted from PubChem (on December 15, 2013) the number of bioassays in which each probe was tested, and the number of bioassays in which the probe was found to be active. Of the 317 chemical-genetic probes, 175 have been tested in bioassays, and they tend to be active in relatively few bioassays (median = 1.4% of the bioassays in which the probe was tested). For those probes for which we have validation data beyond confirmation of drug-induced haploinsufficiency by individual strain analysis, only 3013-0144 appeared active for the majority of the bioassays in which it was tested (5 of 6 bioassays). Interestingly, the probe was active in a bioassay involving the mammalian polo-like kinase PLK-1 as a target, which co-localizes to the cytokinetic furrow, a structure that shares similarity with the yeast septin ring. This is consistent with data in this study supporting the inhibition of septin by 3013-0144. Taken together, these results suggest that the chemical-genetic probes are not promiscuous or broadly reactive.

Gene Ontology (GO)-based enrichment analysis

We obtained Gene Ontology (GO) annotations of yeast genes from SGD (downloaded on May 26, 2012), and additional protein complex annotations were obtained from Benschop *et al.* 2010 (*32*). Unless otherwise indicated, GO biological processes that were too specific (contain less than five genes) or too general (contain greater than 300 genes) were excluded from the analysis. However, we allowed smaller GO categories for protein complexes (>2 genes) and larger categories for cellular components (>300 genes).

Given a query set of genes (e.g. genes with significant FD scores in a HIPHOP profile), we used the hypergeometric test to obtain a *P* value estimating the significance with which the set is enriched with genes annotated to a given GO category. FDR values were obtained from the resulting *P* values using the Benjamini and Hochberg method of adjusting for multiple comparisons (*33*).

For some query sets, the number of GO categories significantly enriched (FDR < 0.1) in the set is large yet many enriched categories are closely related. To facilitate interpretation of the results for a given query, we pruned the set of enriched categories using an adaptation of the GO-Elite method (*34*). Briefly, for enriched categories in the same lineage in the GO hierarchy, we identified the most significantly enriched category and pruned the others. In addition, we assessed the genes that drove the enrichment of each remaining category, i.e. genes in the overlap of the query set and the category set. We clustered the remaining categories based on similarities in their overlap gene sets, measured by the Jaccard coefficient, using the Markov clustering algorithm (MCL, a network clustering algorithm, with inflation = 2). For each cluster, we identified the most significantly enriched category and pruned the others. **Fig. S2** shows the

biological processes enriched among the genes with significantly positive FD scores (standard normal distribution $P < 0.001$) in a fluconazole HIPHOP profile (SGTC_1788), after pruning.

We visualized most GO enrichment results with enrichment maps (**fig. S2**, all enrichment maps available on the website), generated using an approach similar to the Enrichment Map Cytoscape Plugin v1.1 (*35, 36*). In contrast to the plugin, the nodes in each map were clustered with MCL (inflation = 2), using the overlap coefficient computed by the plugin as the similarity metric (coefficients less than 0.5 were set to zero). Nodes in the same cluster were assigned the same node color. Moreover, the size of a node was made to be proportional to the significance with which the corresponding GO category is enriched [$-\log_{10}$(FDR)].

Identification of frequent flier genes and frequent GO terms

If a strain has a significantly positive FD score (standard normal distribution $P < 0.001$) in many HIPHOP profiles, the deleted gene is potentially involved in multi-drug resistance (MDR). Therefore, for each strain, we counted the number of profiles in which the strain has a significantly positive FD score. We defined strains with count ≥ 90 as outliers (90 is the bend point in the positive count distribution, **fig. S20**). We also counted the number of profiles in which a strain has a significantly negative FD score (standard normal distribution $P < 0.001$), and defined outlier strains as those with count ≥ 50 (the bend point in the negative count distribution). There are 287 strains that are outliers in the positive count distribution but not in the negative count distribution and we refer to the genes deleted from these strains as 'frequent flier' or FF genes (**table S9**). These FF genes (~5% of the genome) are enriched for genes with human orthologs (hypergeometric test $P < 10^{-4}$), and include genes involved in drug-related processes such as *PDR1* and *YRR1* encoding regulators of MDR transporters, and *SNQ2* and *PDR16* encoding an ATP-binding cassette (ABC) transporter and phosphatidylinositol transfer protein, respectively, neither of which have previously been identified as MDR genes in fitness assays (*7, 9*). FFs are enriched for genes involved in biological processes known to be associated with MDR, including vesicle-mediated transport and amino acid metabolism (hypergeometric test FDR < 0.1, (*7*)).

In addition to individual FF genes, we identified 'frequent terms' (FTs; **fig. S20**), which represent biological processes, cellular components and protein complexes most frequently affected by compounds. FTs include many processes and components involving FF genes (**table S9**, **fig. S20**). For example, 'protein localization' was affected by most of the compounds screened (~77%) because it includes FF genes such as *RCY1* and *RHO1*, and more specifically, the nuclear pore was perturbed by 27% of the compounds via *GLE2*, *SAC3*, and others. FTs uncovered in this study that were not previously associated with MDR include RNA processing (49% of the compounds, via e.g. *HSH49*, *RPB7*), and protein complex organization/biogenesis (66% of the compounds, via e.g. *TUB3, ARP4*). To identify FTs, we first performed GO enrichment analysis without pruning (see previous section) on the significantly scoring genes in each HIPHOP profile (standard normal distribution $P < 0.001$). For each category, we integrated the 3356 enrichment *P* values (one for each profile) using the weighted-*Z* method [equally weighting all *P*, (*37*)] to obtain a single *Z* score measuring the overall enrichment of the category in the dataset. For each category class, we identified the bend point in the Z distribution; i.e. -110 for biological processes, -70 for cellular components and -260 for protein complexes (**fig. S20**). Categories with *Z* greater than the relevant bend point are considered outliers, and thus, more enriched than other categories in the same class. In a given class, we simplified the list of outlier categories by reducing it to the categories that are highest in the GO hierarchy, and therefore

cover redundant categories in the list. We refer to the remaining categories as frequent terms. The top 10 (i.e. by *Z*) frequent biological processes, cellular components and protein complexes are shown in (**fig. S20**), together with the top 10 processes, components and complexes most frequently identified by HIP hits (where the frequency is the number of compounds with a HIP hit annotated to the given gene category). Since some of the top 10 frequent processes are also in the top 10 processes most frequently hit by HIP hits, **fig. S20** shows less than 10+10=20 processes due to the overlap. The same is true for the top components and complexes.

Identification of enriched chemical moieties

Fragments were generated for all compounds screened using the 'Generate Fragments' component in Pipeline Pilot v8.0. Three different fragment types were generated for each compound. These were 'Murko scaffolds', [contiguous ring systems plus chains that link two or more rings (*38*)], 'Ring Assemblies' (contiguous ring systems), and 'Rings' (individual rings). The number of times each fragment (scaffold, ring or ring assembly) appeared in each category was counted and compared with how often it appeared in the entire dataset. The significance of the count was calculated using the hypergeometric test. FDR values were obtained from the resulting *P* values using the Benjamini and Hochberg method of adjusting for multiple comparisons (*33*).

The definition of cationic amphiphilic drugs (CADs) in Halliwell 1997 (*39*) was used together with Open Babel software (*40*) to identify CADs profiled in our study.

Cofitness and coinhibition

Our chemogenomic dataset can be represented as a matrix *D* where columns are individual HIPHOP profiles and thus represent profiled compounds, and where rows represent deleted genes. By replacing insignificant FD scores (standard normal distribution $P > 0.001$) with 0, resulting in D_{sig}, we can focus on the most significant components of the cellular responses to chemical perturbations identified in this study. In particular, we compute *coinhibition*, defined as the Pearson correlation between two columns in D_{sig}, and a measure of the similarity in the cellular responses to the profiled compounds. We also compute *cofitness*, defined as the Pearson correlation between two rows in D_{sig}, and a measure of the similarity in the importance of the genes in the cellular responses to chemical perturbations. As a result, cofitness can also be considered a measure of functional similarity. We estimated the significance of each correlation using the Student's *t*-distribution method.

To help determine whether relationships between genes, and by extension, relationships between biological processes, supported by cofitness are supported by other types of interactions, we examined whether the genes involved in a significant cofitness interaction (Student's *t*-distribution $P < 0.01$) also exhibit genetic and/or physical interactions reported in BioGRID (version 3.1.91). Cofitness interactions that were not redundant with other types of gene interactions are more likely to be uncharacterized.

Clustering HIPHOP profiles

In identifying major, minor and unique response signatures, we first aimed to identify clusters of HIPHOP profiles that capture similar cellular responses, and we assume that more robust clusters lead to the identification of more robust response signatures. Towards identifying robust clusters, we clustered our HIPHOP profiles using the following methodologies with the indicated variations:

i) hierarchical clustering

Profiles were hierarchically clustered using (1 – coinhibition) as the distance metric, and the Ward agglomeration method. Discrete clusters were derived from the resulting dendrogram using a dynamic branch cutting method (*41*) with the following non-default parameter settings: minimum cluster size = 3, deepSplit = 4, omission of the partitioning around medoids step.

parameter variations: maximum joining height = 35.4 (default), 20, 10, 5, 1

major variant: application of a FD score-based profile filter prior to clustering

ii) Markov clustering [MCL; (*42*)]

Profiles were clustered using coinhibition as the similarity metric, with the negative coinhibition values replaced with zero. (The default inflation value of 2 resulted in a single cluster consisting of all profiles, and thus was not useful for identifying groups of similar profiles.)

parameter variations: inflation = 3, 4, 6, 8, 10

iii) biclustering

Profiles were biclustered with the iterative signature algorithm (*43*), setting the row and column directions to 'upDown' and 'up', respectively. We ran the algorithm using 1000 random seeds generated with sparsity values of 2, 5, 10, and 100. For the normalization step, we computed robust *Z* scores (see The fitness defect (FD) score section). The results were filtered to identify unique and robust biclusters, based on 100 permutations (see the documentation of the isa2 package for R).

parameter variations: standard deviation threshold = 2.3, 3.1

We evaluated each set of clusters by computing silhouette widths, using (1 – coinhibition) as the dissimilarity metric, resulting in a value for each profile quantifying how well-clustered it was (where higher values indicate that profiles were clustered better). For biclusters, where a profile may be a member of multiple biclusters, we adapted the silhouette statistic; when computing the average dissimilarity of a given profile to profiles in other biclusters, we restricted the other biclusters to those that do not contain the given profile. In order to evaluate each individual cluster/bicluster, we computed the median silhouette width across the member profiles (e.g. **table S10**). To facilitate drawing statistically significant conclusions from clusters, we restricted focus to clusters with ≥3 profiles. To avoid drawing overly general conclusions, we also restricted focus to clusters with ≤300 profiles. For each set of clusters, we computed the median of the remaining cluster silhouette widths to serve as a summary statistic for the set (**fig. S21**).

The number of profiles included in clusters varied substantially (from 378 to 3345 profiles), depending on the methodology and parameters used (**fig. S21**). An ideal clustering result strikes a balance between silhouette statistics and profile coverage and this was used to select the best set of clusters from each methodology. With this approach, the best clusters were generated with:

i) hierarchical clustering with dynamic branch cutting: maximum joining height = 35 (the default)
ii) MCL: inflation = 8
iii) biclustering: standard deviation threshold = 2.3

Moreover, of these methodologies, hierarchical clustering generated the best clusters. Under the assumption that stronger signals may result in stronger clusters, we examined the maximum |FD|s of the profiles with silhouette width > 0 given the best hierarchically-derived clusters. Noting that the first quartile of these maximum |FD|s is 10.36, we identified profiles with maximum |FD| ≥ 10 and clustered them hierarchically, then identified discrete clusters using dynamic branch cutting with the default maximum joining height. That is, we applied a FD-based profile filter prior to hierarchical clustering, and we used the resulting clusters in the following analysis.

To assess whether certain clusters are identified by multiple methodologies, and are thus robust, we used the best clusters resulting from hierarchical clustering (without the profile filter) as a reference set. Specifically, we quantified the overlap between each reference cluster and each cluster resulting from a different methodology, using the following metrics (**table S10**):

overlap coefficient = <size of cluster overlap> /
<size of the smaller of the two clusters>
Jaccard coefficient = <size of cluster overlap> /
<size of the union of the two clusters>

For example, when comparing the reference clustering result to that produced by the MCL method, if a reference cluster substantially overlaps with only one MCL cluster, this one-to-one mapping would suggest that the reference cluster is robust to choice of method (**fig. S22**). If instead, there is a one-to-many mapping between a reference cluster and several MCL clusters, the reference cluster was essentially identified by the alternative methodology, but in the form of subclusters; this type of mapping also supports the robustness of the reference cluster, though not as strongly as a one-to-one mapping (**fig. S22**). In addition to determining the mapping from reference clusters to MCL clusters, we also determined the mapping to biclusters, and to clusters resulting from hierarchical clustering with the profile filter. The greater the number of methodologies with a strong mapping (e.g. one-to-one) to a specific reference cluster, the greater the support for the robustness of that cluster. We subsequently assigned a confidence level for each reference cluster according to how strongly the mappings support its robustness, as follows:

i) very strong: one-to-one with Jaccard coefficient ≥ 0.5 for ≥2 methodologies
ii) strong: one-to-one/one-to-many/many-to-one with overlap coefficient ≥ 0.5 for ≥2 methodologies
iii) good: one-to-one/one-to-many/many-to-one with overlap coefficient ≥ 0.5 for ≥1 methodologies
iv) moderate: one-to-one/one-to-many/many-to-one/many-to-many overlap coefficient ≥ 0.5 for ≥2 methodologies
v) other: for all other reference clusters

Forty-eight (80%) of the clusters were assigned a level that is moderate or better, indicating that the majority of the clusters generated by the hierarchical clustering methodology appear robust. If we instead select the ‘best’ clusters based solely on silhouette statistics (i.e. MCL clusters with inflation = 3, biclusters with standard deviation threshold = 3.1), we find that they do not include ≥65% of the 48 high-confidence clusters (overlap coefficient ≥ 0.5), including 2/4 = 50% of the clusters with strong confidence levels. For our chemogenomic dataset, selecting

clusters based on silhouette statistics alone therefore appears too restrictive, whereas finding clusters with relatively high silhouette statistics and profile coverage resulted in the identification of more robust clusters. As a final step towards identifying the most robust clusters, we focused on the 45 high-confidence clusters that each contain >4 profiles, and we refer to the response signatures derived from these clusters as 'major'.

In total, the 45 robust clusters cover 1198 profiles. Most (>99%) of the remaining profiles have ≥1 gene with a significant FD, although they tend to have fewer genes with significant FD, making it more difficult to cluster these profiles robustly (since our profile similarity metric favors profiles with multiple significant genes). However, our hierarchical clustering with maximum joining height = 1 revealed that 1334 of the remaining profiles organize into 334 smaller clusters (median size = 4 profiles, 78% of the clusters include 4 profiles or less). Several of these smaller clusters are of biological interest (we highlight one example in the main text), even though they are of lower confidence statistically because they are supported by fewer profiles. Therefore, we refer to the signatures derived from these clusters as 'minor' (**table S1**). The remaining 824 profiles could not be incorporated into clusters with >2 profiles using our hierarchical clustering methodology. Therefore, we assume these profiles define unique signatures.

Defining response signatures

For each cluster, we computed a median profile, where for each deletion strain, we assigned the median FD score across the profiles in the cluster. Strains with significantly positive scores (standard normal distribution $P < 0.001$) in the median profile (i.e. by definition, exhibit significant fitness defects in at least 50% of the profiles in the cluster) identify genes that are particularly important to the responses captured by the profiles in the cluster, and we refer to this significant subset of the median profile as the *response signature*. We performed GO enrichment analysis on each response signature.

Each response signature is associated with a set of HIPHOP profiles, and by extension, a set of compounds. We therefore tested for structural enrichment in each signature by testing its compound set for significant enrichment of chemical moieties, relative to all profiled compounds. Significantly enriched chemical moieties (hypergeometric test FDR < 0.1) that also comprise ≥20% of a compound set are shown in **Fig. 2**.

The response signatures shown in **Fig. 2** were manually named based on the functional or structural characteristics of the signatures.

Identification of enriched bioactivity classes

To determine whether the set of compounds associated with a given signature is significantly enriched for compounds known to have similar mechanisms, we used bioactivity classes obtained from PubChem (on August 15, 2013). These bioactivity classes specify compounds annotated to the same class by ChEBI, KEGG, MeSH, LIPID MAPS or the International Patent Classification 2013. Compounds annotated to the same class are typically known to have similar bioactivity (e.g. therapeutic use), and in some cases, similar mechanism. We restricted our focus to classes with >1 compound profiled in this study, and to signatures associated with >1 compound annotated to a bioactivity class. For each remaining set of signature compounds, we used the hypergeometric test to obtain a *P* value estimating the significance with which the set is enriched with compounds annotated to a given bioactivity

class. FDR values were obtained from the resulting *P* values using the Benjamini and Hochberg method of adjusting for multiple comparisons (*33*). See **table S5**.

Identification of response signatures in genetic interaction profiles

Using yeast large-scale genetic interaction profiles (*18*), we identified the set of genes that genetically interact with each genetic interaction screen query gene (using the intermediate cutoff). Each set of interaction partners was assessed for significant GO enrichment as described above (biological processes, molecular functions, protein complex). We found that 380 query genes do not have interaction partners that are significantly enriched for any GO category (hypergeometric test FDR < 0.1). For each of these query genes, we used the hypergeometric test to obtain a *P* value estimating the significance with which its negative interaction partners are enriched for each major response signature (with >2 genes).

Detection of response signatures in another chemogenomic dataset

To determine whether the major response signatures are similar to signatures obtainable from the Hillenmeyer *et al.* 2008 (*7*) dataset using the same approach, we restricted our focus to compounds that were profiled by both HIP and HOP in that study. Since some compounds were profiled at different concentrations, we randomly sampled one concentration profile for each compound and clustered the resulting set of profiles using hierarchical clustering as was done to define the major response signatures, and repeated 50,000 times. To define an overall set of clusters from the sample-based clusters, we did the following:

i) for each pair of compounds, compute the percentage of samples where the compounds ended up in the same cluster
ii) use (1 – <same cluster percentage>) as a distance metric to cluster the compounds with hierarchical clustering as before

To identify the robust clusters in this set, we applied the FD-based profile filtered described above, and defined another overall set of clusters given the restricted set of profiles for sampling. Performing the overlap-based mapping described above with these two sets of clusters (derived with and without the profile filter), we identified 26 robust clusters. With these clusters, we defined Hillenmeyer response signatures. The similarity between signatures from this study and Hillenmeyer signatures was measured using the overlap coefficient (<size of overlap between signature gene sets> / <size of the smaller of the two signature gene sets>). If a signature from this study has overlap ≥ 0.5 with some Hillenmeyer signature, we claim that the signature from this study was detected in another chemogenomic dataset.

Identification of response signatures using subsets of HIPHOP profiles (saturation analysis)

To determine whether the major response signatures can be derived from subsets of the HIPHOP profiles, we repeatedly sampled random subsets (of increasing size) of our HIPHOP profiles, hierarchically clustered them as was done to define the major response signatures, and defined new signatures from these clusters. The similarity between major signatures and signatures derived from subsets was measuring using the overlap coefficient (<size of overlap between signature gene sets>/<size of the smaller of the two signature gene sets>). If a major signature has overlap ≥ 0.5 with some subset-derived signature, we claim that the major signature was identified by the subset of profiles. This process was repeated for 100 samples for each subset size.

Drug-induced Phospholipidosis (DIPL) prediction

A naïve Bayes structure model using ECFP_4 fingerprints (*44*) was created to predict whether or not a given compound induces the *NEO1* or *NEO1-PIK1* response signature. *NEO1* and *NEO1-PIK1* compounds were considered active, while all other compounds screened were considered inactive. The charge and stereo form of each compound was standardized, and salts were removed using Pipeline Pilot version 6.1 (Scitegic Inc. Accelyrs, San Diego, CA, USA). The robustness of the model was tested using a five-fold cross validation procedure, which was repeated five times. In each cross-validation cycle, $4/5^{th}$ of the data are used to build the model (i.e. the training set), and then the performance of the model is benchmarked on the remaining fifth (i.e. the test set). This procedure is repeated five times so that all molecules are in the test set once and in the training set four times in one cross-validation procedure. The performance of the model is then reported as the average enrichment factor (i.e. <the number of the actives found> / <number of expected actives found by random sampling>) from all of the cross-validation runs for the top 5%, top 10%, top 20% and top 50% after ranking compounds by the model score (**fig. S16**).

The performance of this model in predicting DIPL was tested by ranking a set of 519 compounds experimentally tested for DIPL, i.e. annotated to the target 'Phospholipidosis' (CHEMBL1626541) in the ChEMBL database (*45*). Of those compounds, 111 were active (shown to result in DIPL in at least one publication in ChEMBL). Fifty-five compounds were removed from the initial set obtained from ChEMBL to avoid potential bias, as they were screened in this study.

Pma1: membrane ATPase assay and media acidification assay

Plasma membranes were prepared as described (*46*) from *Saccharomyces cerevisiae* Σ 1278b strain. ATPase assays were performed in a 96-well microplate at 30ºC in reaction medium (4mM MgATP, 1mM free Mg^{2+} ($MgCl_2$), 50mM Mes (pH 6.5), 20mM KNO_3, 0.2mM $NaMoO_4$, 0.05% Brij, and 10 mM sodium azide). The reaction was started by adding 10µL of 4.34mM ATP (the total amounts of ATP and Mg^{2+} to be added to the reaction mixture at the indicated pH were calculated in order to obtain 4 mM MgATP with 1mM free Mg^{2+}) to 40µL of reaction medium containing 1.5µg of protein in the presence or absence of 0.2mM orthovanadate. After 9min, 60µL of 1% SDS was added to stop the reaction, and 30µL of 50% (w/v) $(NH_4)_6Mo_7O_{24}$, 4 N H_2SO_4 and 30 µL of 1% 4-methylaminophenol sulfate, 3% $NaHSO_3$ were added successively. The OD was finally measured at 700ηm.

For the whole-cell media acidification assay, a saturated yeast culture (BY4743) was diluted to 0.25OD/mL in 40mL of YPD. The culture was grown for 3h at 30ºC and then 10mL aliquots were pelleted at 3000rpm for 5 minutes, washed twice with 40mL of water, resuspended in 40mL water, and incubated on ice for 2h. Cells were then collected at 3000rpm for 5 minutes. The supernatant was removed and the cell pellet was resuspended in 15mL of 25mM KCl. DMSO or chemical inhibitor was added to the cell suspension and the pH of the solution was monitored with a pH probe (Beckman), while mixing using a magnetic stir bar. The pH was allowed to stabilize for ~5 minutes and then 750µL of 40% glucose was added and the pH was recorded every 30 seconds for 10 minutes.

Sec14: phosphatidylinositol (PtdIns) transfer assays

Recombinant His_8-Sec14 protein was expressed in *E.coli* (BL21 DE3; Novagen), purified by TALON metal affinity resin chromatography (Clontech), and stored in 300mM NaCl, 25mM

Na_2HPO_4 (pH=7.5), 5mM 2-mercapthoethanol (*47, 48*). Protein concentration was determined by Coomassie Brilliant Blue staining after SDS-PAGE and comparison to BSA standards. The purity of Sec14 protein was typically >95%.

PtdIns transfer assays measured the mobilization of radiolabeled phospholipid from rat liver microsomes to liposomes, and from liposomes to bovine heart mitochondria as described previously (*49, 50*). In experiments involving chemical inhibitors, purified recombinant Sec14 was pre-incubated in the presence of acceptor membranes and compound for 30 minutes at 37ºC prior to initiation of assay by addition of radiolabeled donor membranes.

Tubulin polymerization, binding-site and imaging assays

Tubulin polymerization was assessed using a fluorescence-based assay according to the manufacturer's instructions (Cytoskeleton, Inc.). Briefly, purified porcine neuronal tubulin was incubated with DMSO (negative control, in duplicate), with 5µM nocodazole (positive control), or with increasing concentrations (5µM, 10µM, and 20µM) of 1327-0036. Polymerization was measured by excitation at 360µm and emission at 420µm in a 96-well plate using a Tecan Genios plate reader.

The tubulin-binding site was determined, as previously described (*51*). Briefly, purified bovine tubulin (5µM) was incubated with 1327-0036 (100µM), vinblastine (100µM) or DMSO at 37°C for 30 minutes. Colchicine (10µM) was added and incubated for an additional 60 minutes and the fluorescence of the colchicine-tubulin complex was then measured at excitation and emission wavelengths of 360nm and 430nm, respectively.

Yeast microtubules were imaged using the diploid *S. cerevisiae* strain (NTY5) with the following genotype: mat a/α, ura3-52::GFP-*TUB1-URA3*/ura3Δ0, leu2Δ0/leu2-801, his3 Δ1/his3-200, in which a tagged copy of alpha tubulin is integrated at the deleted URA3 locus. Cells were grown in YPD until they reached an OD_{600} of 0.5, compound or vehicle control (DMSO) was added and cells were incubated for 2h at 30ºC prior to fixation by addition of methanol-free formaldehyde to 1% final concentration. Fixed cells were adhered to poly-lysine coated coverslips, assembled as wet mounts on a microscope slide and imaged on a Ziess Axiovert 200 using a 40X objective. Phenotypes were assessed as normal or abnormal for 100 cells in each condition at the following stages of the cell cycle: unbudded (G1), small bud (S), medium bud (metaphase), large bud (anaphase).

HeLa cells containing an integrated copy of GFP-labelled alpha tubulin under the control of the constitutive CMV promoter (*52*), were plated onto glass coverslips and grown in Dulbecco's Modified Eagle Medium containing 10% fetal calf serum at 37ºC in a 5% CO_2 atmosphere. Cells were treated with DMSO (0.05% final concentration) and nocodazole or 1327-0036 and returned to the incubator for 120 minutes. After additional staining with Hoechst 33342 dye for 30 minutes, cells on coverslips were inverted onto a microscope slide and sealed with VALAP (vaseline, paraffin and lanolin at 1:1:1 ratio) leaving 2mm openings on each side of the coverslip to permit media exchange. Cells were imaged on a Quorum WaveFX spinning disc confocal microscope system using a 40X objective. All images were collected at the same exposure times.

Actin polymerization and imaging assays

Polymerization of pyrene-conjugated rabbit skeletal muscle actin was measured using the Actin polymerization Biochem Kit (Cytoskeleton, Inc.). Briefly, G-Actin was centrifuged at

40,000xg for 2h and the supernatant was carefully transferred to a new tube without touching the pellet. 1.8mL of the solution containing G-actin monomers were mixed with 3.6μL ATP. 200μL of this mixture were added to the wells of a microtiter plate and polymerization of the filaments was followed (in the presence or absence of chemical inhibitor) by measuring fluorescence at 415ηm in a Sunrise plate reader (Tecan). Arp2/3-dependent actin polymerization was performed essentially as above, only purified Arp2/3 complex and the VCA domain of WASP was included in the reaction (Cytoskeleton, Inc.).

Actin labelling of yeast cells with fluorescent phalloidin was performed essentially as described previously (*53*). Briefly, cells were grown to early log phase and fixed by the direct addition of formaldehyde (3.7% final) to the medium buffered at pH 6.8. After a 2h fixation, the fixative was removed by washing two times with water and two times with PBS set to pH 7.4. The cells were permeabilized for 10 minutes using 0.2% Triton X-100 in PBS. After two washes in PBS, the cells were incubated for 2h with 1.4μM (final concentration) Acti-stain™ 488 Fluorescent Phalloidin (Cytoskeleton, Inc.) in the dark at 25°C. After two washes in PBS, the cell pellet was resuspended in 20μL of PBS and then applied to poly-lysine coated coverslips. For differential interference contrast (DIC) and fluorescence microscopy, a system with the following components was used: Axiovert 40 CFL (microscope from Zeiss), AxioCam MRm (camera from Zeiss), LD A-Plan 40x, 1006-595 (Ph2) (objective from Zeiss), Axiovision Rel. 4.5 (software from Zeiss), X-Cite Series 120 (lamp from EXFO). To visualize the polymerized actin filaments (i.e. the products of the *in vitro* reaction), a fresh dilution of Acti-stain™ 488 Fluorescent Phalloidin was prepared by diluting 2μL in 98μL water (1.4μM final concentration). The staining solution was mixed 1:1 with the polymerized actin filament samples, applied to a poly-lysine coated coverslips, and imaged as above.

Mammalian cell actin morphology assays were performed as follows; A549 cells were grown in RPMI with 10% serum to 80% confluency on acid washed coverslips and fixed by the direct addition of formaldehyde (1.0% final concentration) to the medium. After 2h, the fixative was removed by washing two times with PBS. Fixed cells were permeabilized for 10 minutes using 0.02% Triton X-100 in PBS. After two washes in PBS, cells were incubated for 2h with 1.4μM (final concentration) Acti-stain™ 488 Fluorescent Phalloidin (Cytoskeleton, Inc.) in the dark at 25°C. After two washes in PBS, coverslips were mounted in media containing Hoechst 33342 dye and imaged on a Quorum WaveFX spinning disc confocal microscope system using a 40X objective. All images were collected at the same exposure times.

Cdc12: Golgi re-orientation assay

HeLa cells where treated with either DMSO, 3013-0144 or 1-(2-chloropyridin-4-yl)-3-phenylurea (forchlorfenuron or FCF). Confluent monolayers of cells were serum starved for 18h and pre-treated with compounds for 1h before a “scratch wound” was introduced. Serum was then added to 10% and the cells were fixed 45 minutes later. Cells were then stained with DAPI to visualize nuclei and anti-GM130 to determine whether the cells had successfully reoriented their Golgi complex towards the wound edge (shown as a white line **Fig. 1B**). The Golgi complex reoriented to the leading quadrant in control cells treated with the DMSO vehicle but, in contrast, treatment with 3013-0144 inhibited Golgi reorientation as effectively as FCF.

Trichlorophene assays

Cell culture. TEX human leukemia cells (a gift from Dr. J. Dick, Toronto, Canada) were maintained in IMDM (Iscove's modified Dulbecco's medium), 15% FBS, 100µg/mL penicillin, 100µg/mL streptomycin, 20ng/mL SCF, 2ng/mL IL-3 and 2mM L-glutamine. OCI-AML2 cells were grown in IMDM medium with 10% FCS, 100µg/mL penicillin and 100µg/mL streptomycin (all from Hyclone, Logan, UT). Cells were incubated at 37^{o}C in a humidified air atmosphere supplemented with 5% CO_2.

Cell growth and viability assay. Cell growth and viability were measured with the CellTiter-Fluor assay (Promega, Madison, WI, USA) as per manufacturer's instructions and by trypan blue staining.

Enzymatic activity of the respiratory chain complexes. Complex activity was analyzed from mitochondria–enriched pellet, prepared by cell lysis and centrifugation, followed by freezing at –80°C for 7 days. Frozen mitochondrial pellets isolated from treated cells were re-suspended in 10mM Tris, 250nM sucrose, 1mM EGTA containing protease inhibitors and Dodecyl–D–maltoside, and analyzed for protein concentration and citrate synthase activity. Complex I activity was detected by monitoring rotenone-sensitive 2,6-dichloroindophenol reduction by electrons accepted from decylubiquinol reduced after oxidation of NADH by complex I (*54*). Complex III activity was assessed with a modified method described by (*55*) and (*56*). Complex III-specific activity was measured by following the increase in absorbance due to the reduction of ferrocytochrome c at 550nm, with 580nm as the reference wavelength. Complex IV activity was measured by KCN-sensitive oxidation of ferrocytochrome c at 550nm, with 540nm as the reference wavelength (*57*). Ferrocytochrome c was prepared by reducing cytochrome c with sodium ascorbate followed by dialysis for 24h (*58*). Citrate synthase activity was measured based on the chemical coupling of CoASH, released from acetyl-CoA during the enzymatic synthesis of citrate to DTNB [Ellman's reagent, 5,5'-dithiobis(2-nitrobenzoic acid)], and the release of the absorbing mercaptide ion was monitored at 412 nm (*59*). The enzyme activity of Complexes I, III, and IV was normalized to citrate synthase activity.

Oxygen consumption and reserve capacity. The rate of oxygen consumption was measured using a Seahorse XF96 analyzer as per manufacturer's instructions (Seahorse Bioscience, North Billerica, MA, USA). After a 72h treatment, cells were washed, re-suspended with un-buffered medium αMEM medium containing 2% FCS for 45min at 37^{o}C in a CO_2-free incubator, and seeded at 120 000 cells/well in XF96 plates. The rate of oxygen consumption (OCR) was measured over time. Spare reserve capacity of the mitochondrial respiratory chain was measured by treating cells in succession with 2µM oligomycin, and 1µM and 1.5µM FCCP (carbonyl cyanide p-trifluoromethoxyphenylhydrazone) for TEX and OCI-AML2, respectively. Oxygen consumption was measured as described above. After the assay, the cell number was quantified with the CyQUANT Cell Proliferation Assay Kit (Molecular Probes, Eugene, OR, USA).

Reactive oxygen species. Mitochondrial reactive oxygen species (ROS) were detected by staining cells with MitoSOX Red (Molecular Probes, Invitrogen, Eugene, OR, USA) (final concentration 5 µM) and flow cytometric analysis as per manufacturer's instructions. Data were analyzed with FlowJo version 7.7.1 (TreeStar).

All experiments were performed in triplicate.

Sec13: CPY processing, Gap1 Activity and localization

Sec13-dependent protein trafficking was assessed by carboxypeptidase Y (CPY) processing. Wild-type cells (ESY360) were grown in synthetic medium lacking methionine, treated with DMSO or 0.5mM (4065-0411 or 4544-0069) for 30 minutes at 24^{o}C, and then pulse-labeled for 10 minutes with ^{35}S-methionine/cysteine. A chase was initiated by addition of cold methionine and cysteine (2mM final), and aliquots were removed at 15, 30 and 60 minutes and added directly to trichloroacetic acid (10% final). CPY was immunoprecipitated and processed as described (*60*).

Wild-type cells (ESY360) were grown in synthetic dextrose (SD) ammonium medium at 24^{o}C, treated with DMSO, 4065-0411 or 4544-0069 for 30 minutes, and then Gap1 activity was assessed by measuring the rate of uptake of [^{14}C]-citrulline as described (*61*). Data shown summarize three independent experiments. Gap1-GFP expressing cells (CKY834) were grown in SD ammonium, labeled with 3µM FM4-64 for 1h, collected by centrifugation and resuspended in fresh SD ammonium medium for 45 minutes to allow FM4-64 to reach the endosome and vacuole membranes. Cells were then treated with DMSO, 4065-0411 or 4544-0069 for 30 minutes prior to Gap1-GFP visualization as described (*62*).

Erg7: LC/MS/MS analysis of 2,3-oxidosqualene and ergosterol content

Wild-type BY4743 yeast cells were grown for 80 generations (via serial dilution every five generations) in YPD 20mM HEPES pH7 with 2% DMSO, or with 60µM 4296-0685 (i.e. IC_{50} based on growth in YPD). 25 OD_{600} units (~5 x 10^8 cells) were harvested by centrifugation, washed 2X in cold PBS, re-adjusted to exactly 20 OD_{600} units (~4 x10^8 cells) and pelleted in a microfuge. The cell pellets were drop-frozen in liquid nitrogen. Compound treatment was performed in technical triplicate, using three independent biological replicates.

For lipid extraction, the cell pellets were resuspended in 300µL PBS and added to a glass tube containing 3mL of a 1:1 dichloromethane-methanol (DCM-MeOH) (v/v) mixture spiked with 0.1µg of the internal standard cholesterol-d_7 (C/D/N isotopes, Pointe-Claire, QC, Canada). The samples were vortexed and centrifuged at 3500rpm for 5 minutes at 22°C. The supernatant was transferred to a new glass tube and the pellet was re-extracted with 3mL of DCM-MeOH. The two extracts were pooled and phase separation was initiated by adding 3mL PBS, after which the samples were vortexed and centrifuged. The lower organic phase was transferred to a new glass tube and the remaining aqueous phase was re-extracted with 3mL of dichloromethane. Both organic phases were pooled and subsequently dried under N_2. These lipid extracts were dissolved in 400µL methanol and analyzed using a 1200 series HPLC coupled to a 6410 Triple Quad LC/MS mass spectrometer (Agilent Technologies) with the ESI source in positive ion mode. Samples were loaded onto a Zorbax XDB-C_{18} column (4.6 x 50mm, 3.5µm) equipped with a C_{18} guard column at 40°C using a flow rate of 0.4mL/min. The mobile phase consisted of HPLC grade water (A) and methanol (B) both supplemented with 5mM ammonium acetate. Lipids were resolved using the following gradient: 0-1min, 96% B; 1-3.3min, 96% to 100% B; 3.3-14.8min 100% B; followed by a 5min re-equilibration time. The MS was operated in multiple reaction monitoring (MRM) mode. The ESI source conditions were as follows: gas temperature 175°C, nebulizer pressure 35psi, drying gas (nitrogen) 10L/min, capillary voltage 6000V. Purified calibration standards ergosterol and 2,3-oxidosqualene were purchased from Sigma Aldrich.

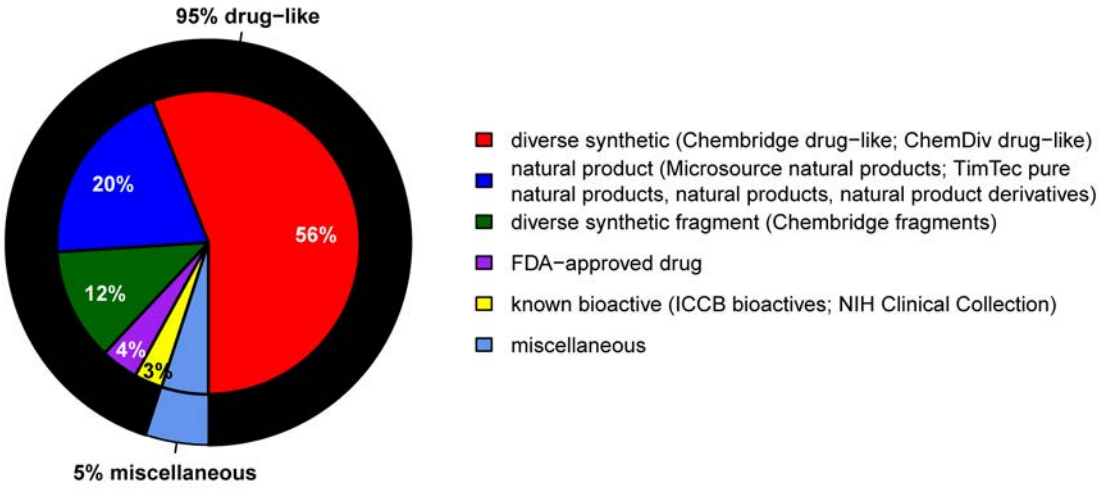


Fig. S1

Molecular categories and sources of compounds screened in HIPHOP. The largest category of molecules came from two diverse synthetic libraries (red) that were selected for maximal structural diversity, from Chembridge (http://www.chembridge.com) and ChemDiv (http://us.chemdiv.com/) commercially available collections. Selected compounds were further filtered by the vendors to avoid highly reactive and unstable compounds (*63*). The Chembridge fragment library (green) comprises low molecular weight compounds (MW ≤ 300, H-bond donors ≤ 3, H-bond acceptors ≤ 3, cLogP ≤ 3). Several small natural product libraries (blue) were obtained from TimTec, e.g. (http://www.timtec.net/Natural-Compound-Library.html), in addition to a collection from Microsource (http://www.msdiscovery.com/natprod.html). FDA-approved drugs (purple) were obtained at the highest purity available from commercial sources. The NIH clinical collection (http://www.nihclinicalcollection.com/) is made up of small molecules that have a history of use in human clinical trials. Similarly, The ICCB bioactive library (http://www.enzolifesciences.com/BML-2840/iccb-known-bioactives-library/) contains 480 biologically active molecules assembled from a variety of sources (yellow). The miscellaneous compounds include well-characterized chemical probes/perturbants ordered from a variety of chemical suppliers (light blue). Because the commercial libraries that we screened use different criteria for 'rapid elimination of swill' (REOS) filters (i.e. elimination of reactive compounds)(*63*), we caution that, as a first pass screen, it must be recognized that the majority of our compounds would require medicinal chemistry and far more characterization to qualify as chemical probes.

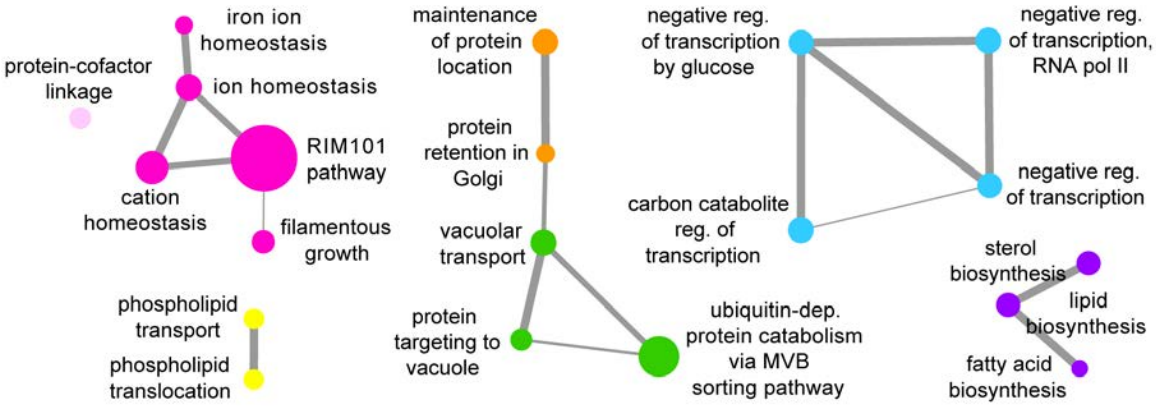


Fig. S2

Gene Ontology (GO) enrichment map for fluconazole. The map indicates GO biological processes enriched in the fluconazole HIPHOP profile shown in **Fig. 1A**. Nodes: enriched biological processes (hypergeometric test FDR ≤ 0.1); size is proportional to the degree of enrichment [$-\log_{10}$(FDR)]. Node color indicates processes that share genes. Edges indicate ≥ 50% gene overlap between processes; width is proportional to degree of overlap.

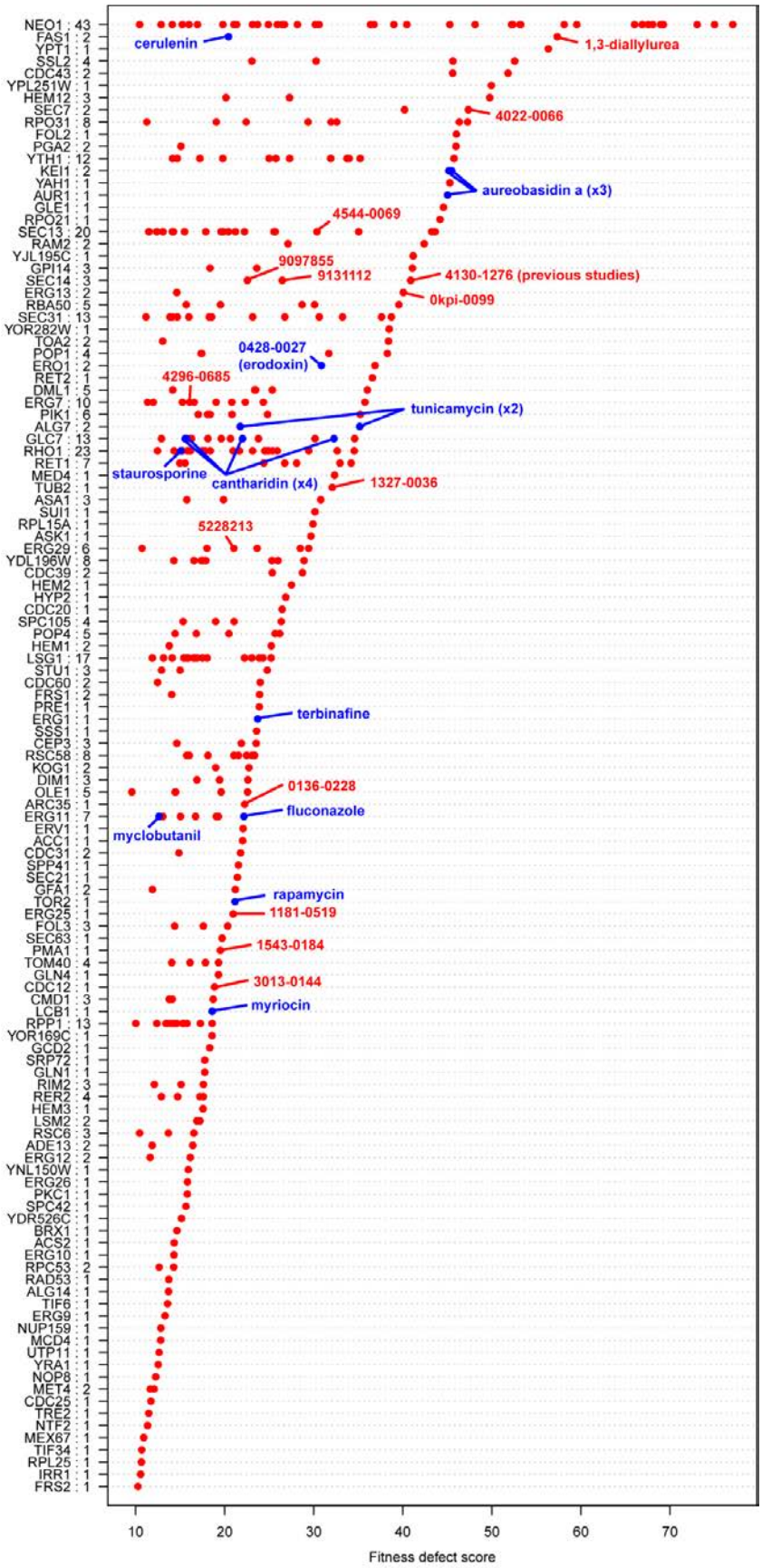
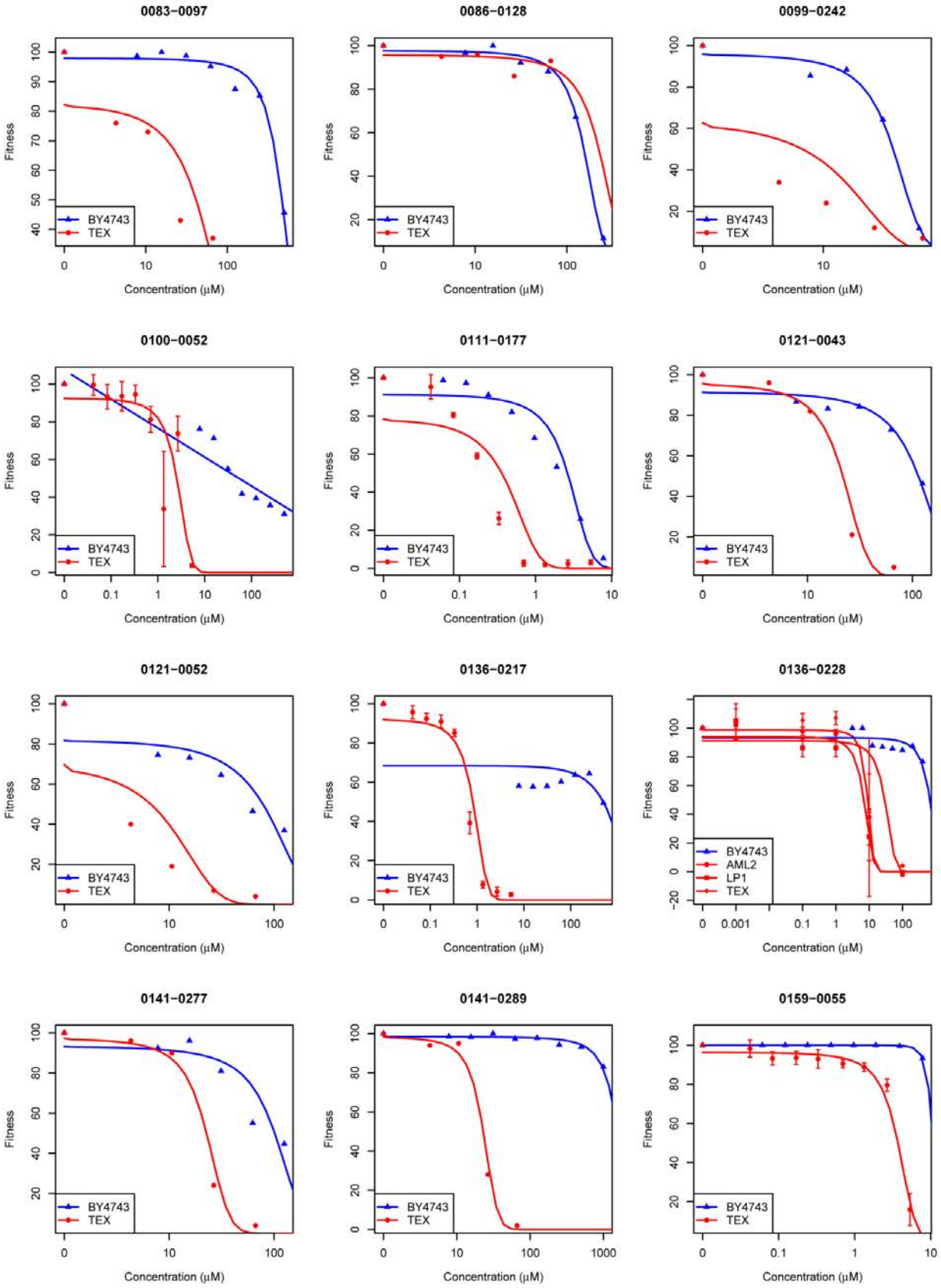
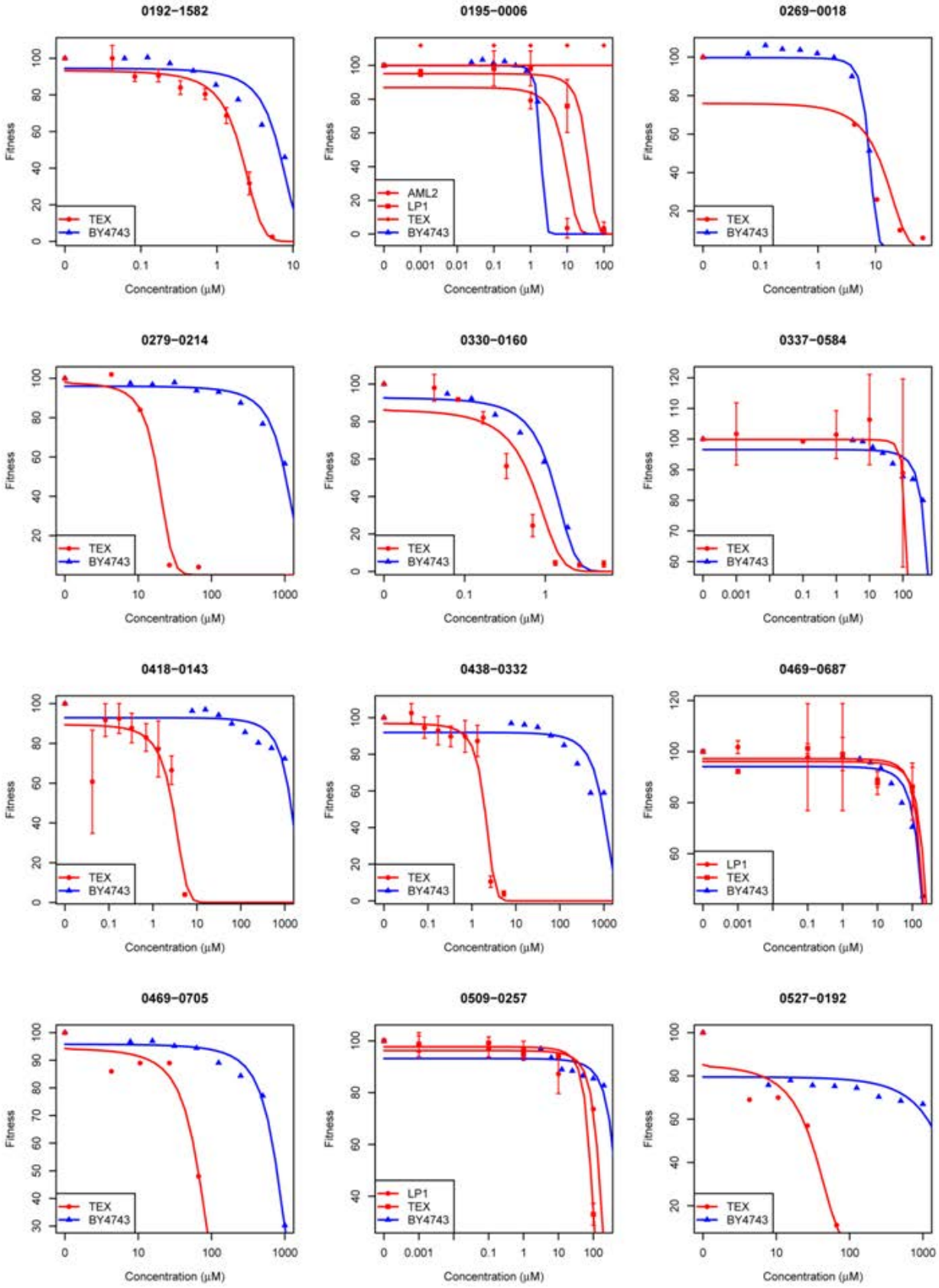
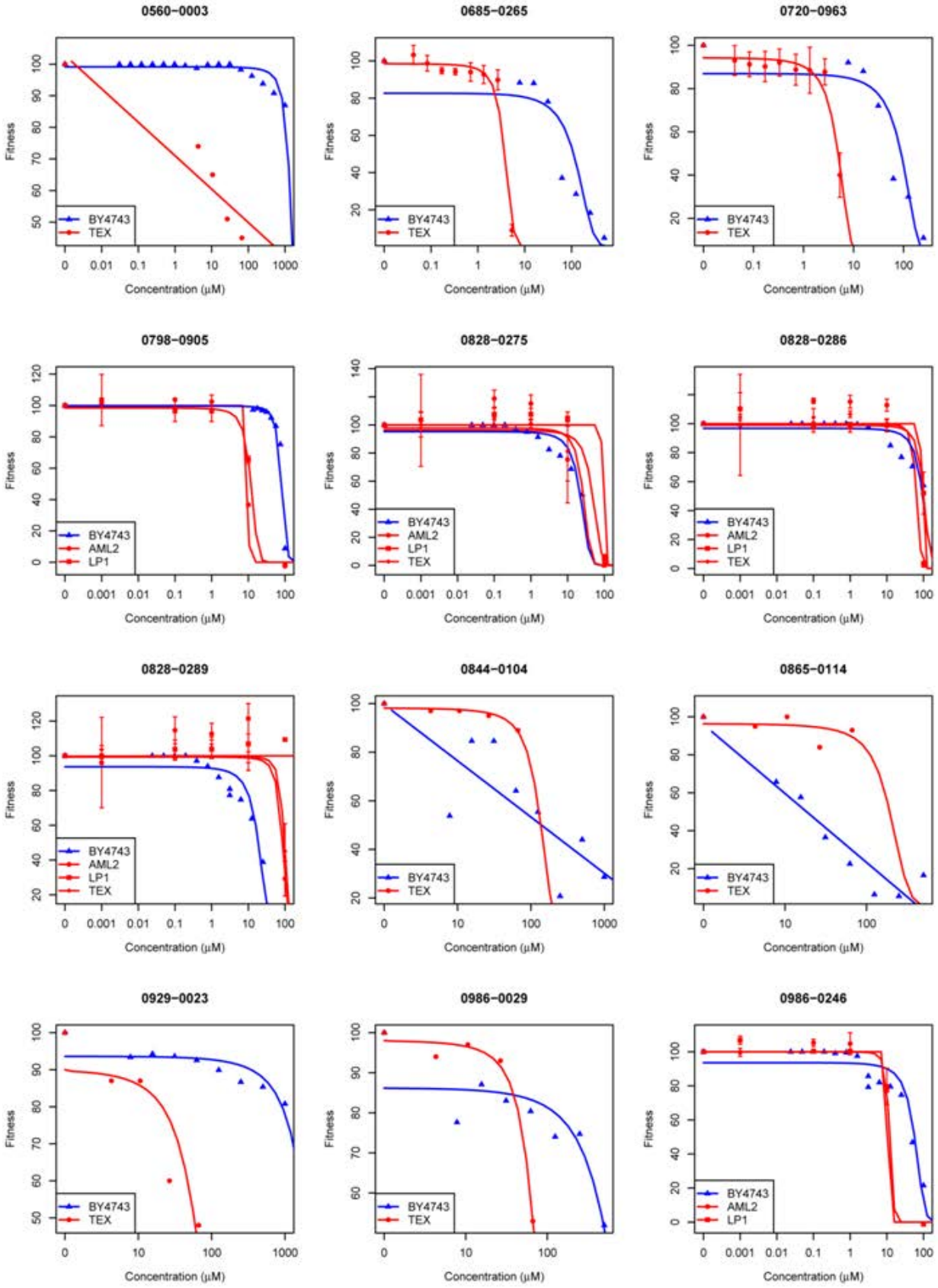

NEO1 : 43
FAS1 : 2
YPT1 : 1
SSL2 : 4
CDC43 : 2
YPL251W : 1
HEM12 : 3
SEC7 : 2
RPO31 : 8
FOL2 : 1
PGA2 : 2
YTH1 : 12
KEI1 : 2
YAH1 : 1
AUR1 : 1
GLE1 : 1
RPO21 : 1
SEC13 : 20
RAM2 : 2
YJL195C : 1
GPI14 : 3
SEC14 : 3
ERG13 : 2
RBA50 : 5
SEC31 : 13
YOR282W : 1
TOA2 : 2
POP1 : 4
ERO1 : 2
RET2 : 1
DML1 : 5
ERG7 : 10
PIK1 : 6
ALG7 : 2
GLC7 : 13
RHO1 : 23
RET1 : 7
MED4 : 1
TUB2 : 1
ASA1 : 3
SUI1 : 1
RPL15A : 1
ASK1 : 1
ERG29 : 6
YDL196W : 8
CDC39 : 2
HEM2 : 1
HYP2 : 1
CDC20 : 1
SPC105 : 4
POP4 : 5
HEM1 : 2
LSG1 : 17
STU1 : 3
CDC60 : 2
FRS1 : 2
PRE1 : 1
ERG1 : 1
SSS1 : 1
CEP3 : 3
RSC58 : 8
KOG1 : 2
DIM1 : 3
OLE1 : 5
ARC35 : 1
ERG11 : 7
ERV1 : 1
ACC1 : 1
CDC31 : 2
SPP41 : 1
SEC21 : 1
GFA1 : 2
TOR2 : 1
ERG25 : 1
FOL3 : 3
SEC63 : 1
PMA1 : 1
TOM40 : 4
GLN4 : 1
CDC12 : 1
CMD1 : 3
LCB1 : 1
RPP1 : 13
YOR169C : 1
GCD2 : 1
SRP72 : 1
GLN1 : 1
RIM2 : 3
RER2 : 4
HEM3 : 1
LSM2 : 2
RSC6 : 3
ADE13 : 2
ERG12 : 2
YNL150W : 1
ERG26 : 1
PKC1 : 1
SPC42 : 1
YDR526C : 1
BRX1 : 1
ACS2 : 1
ERG10 : 1
RPC53 : 2
RAD53 : 1
ALG14 : 1
TIF6 : 1
ERG9 : 1
NUP159 : 1
MCD4 : 1
UTP11 : 1
YRA1 : 1
NOP8 : 1
MET4 : 2
CDC25 : 1
TRE2 : 1
NTF2 : 1
MEX67 : 1
TIF34 : 1
RPL25 : 1
IRR1 : 1
FRS2 : 1
cerulenin
1,3-diallylurea
4022-0066
aureobasidin a (x3)
4544-0069
9097855
9131112
4130-1276 (previous studies)
0kpi-0099
0428-0027
(erodoxin)
4296-0685
tunicamycin (x2)
staurosporine
cantharidin (x4)
1327-0036
5228213
terbinafine
0136-0228
fluconazole
myclobutanil
rapamycin
1181-0519
1543-0184
3013-0144
myriocin
10
20
30
40
50
60
70
Fitness defect score

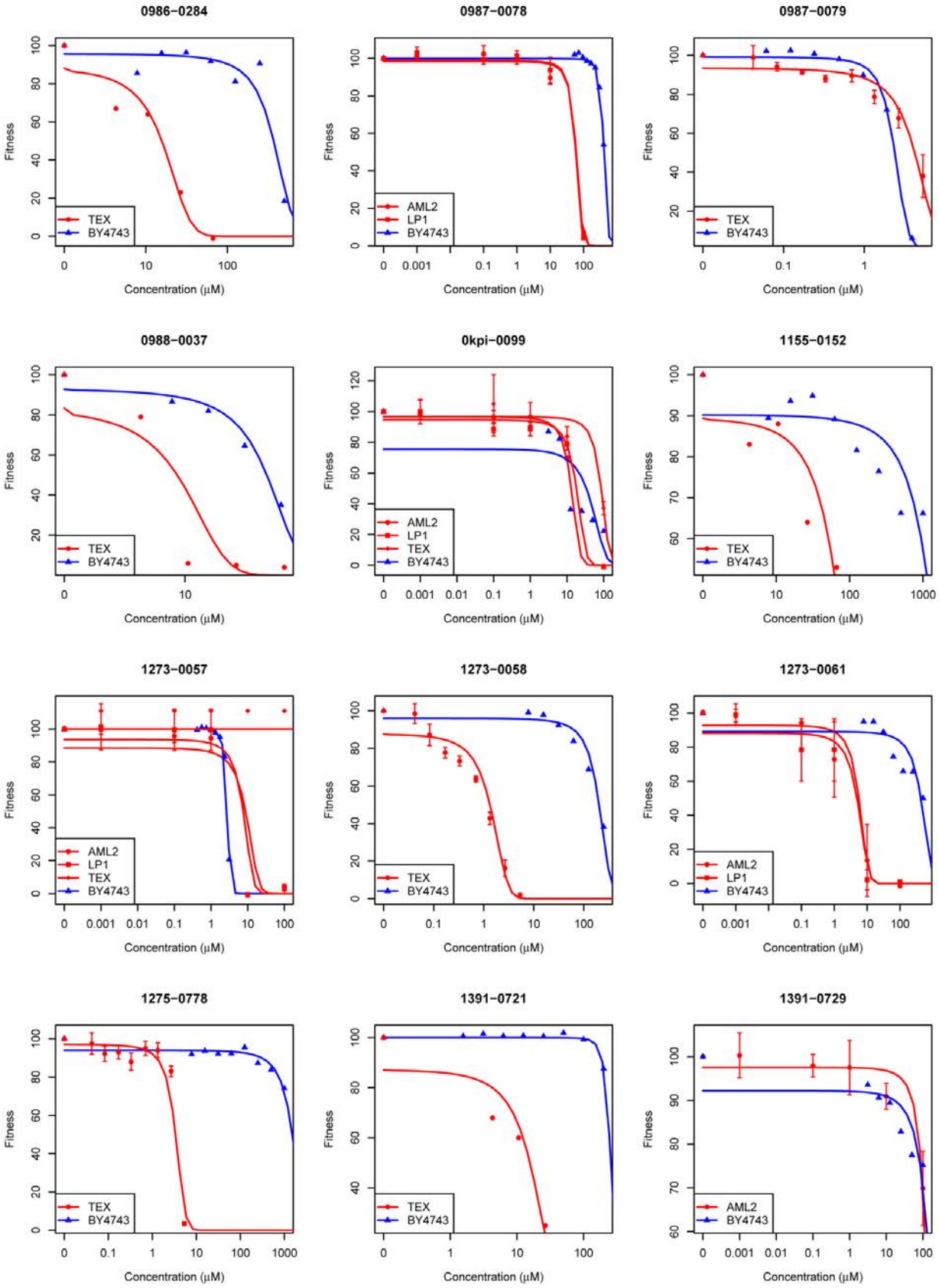
Fig. S3 (previous page)

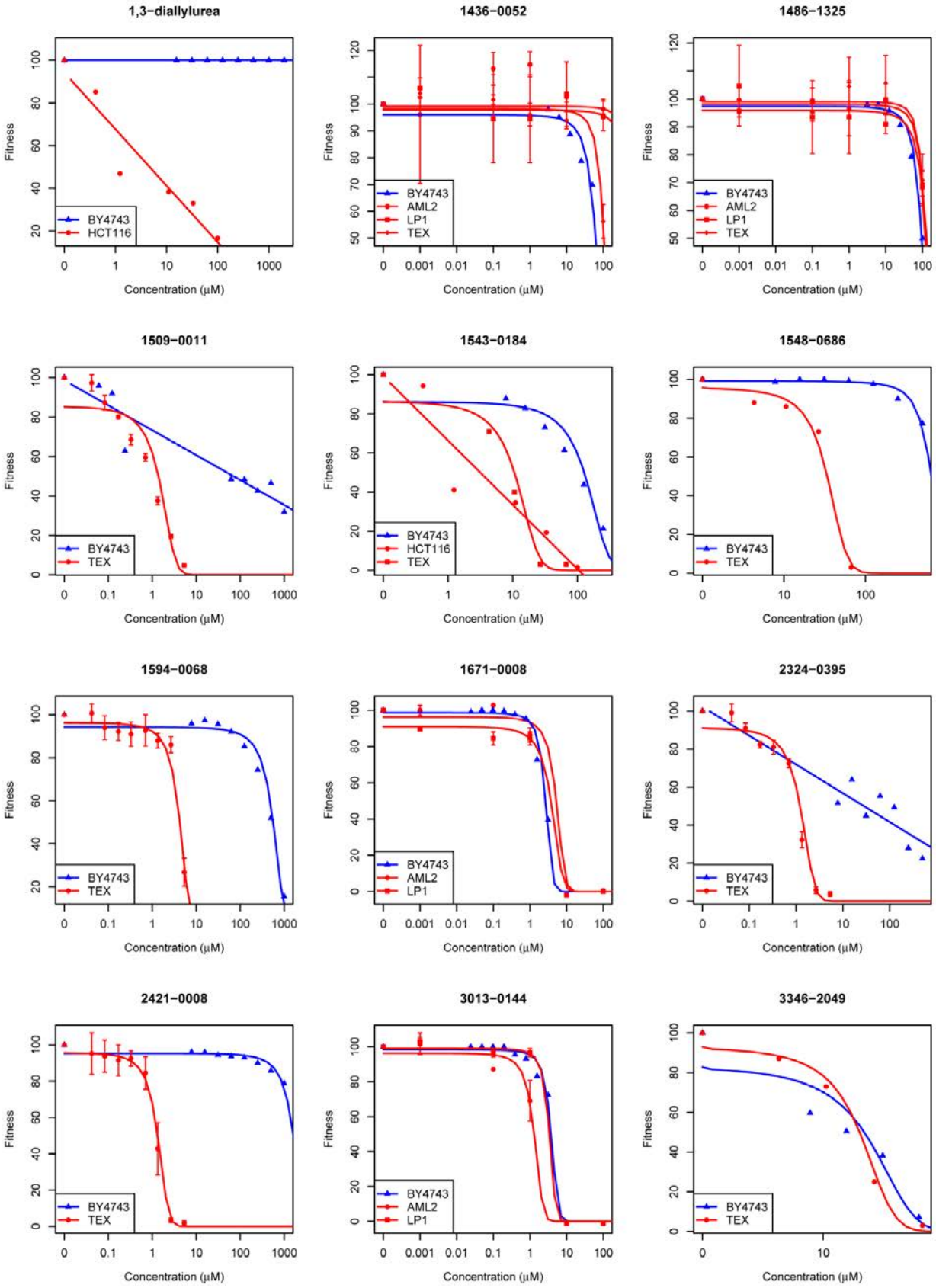
Fitness defect scores of HIP hits. Genes identified as HIP hits, by the criteria described (*10*), are indicated by name (y-axis) followed by the number of times they were identified across all HIP profiles (value following the gene name). For each gene, we plotted its significant fitness defect scores (x-axis, standard normal distribution $P < 0.001$) from profiles where the gene was identified as a HIP hit; thus, each point represents a putative compound-target interaction between the profiled compound and the gene's protein product. Established compound-target interactions are indicated in blue and uncharacterized interactions, supported by results from additional *in vivo* and/or *in vitro* assays, are labeled in red. HIP hit genes are plotted in descending order by maximum fitness defect score.

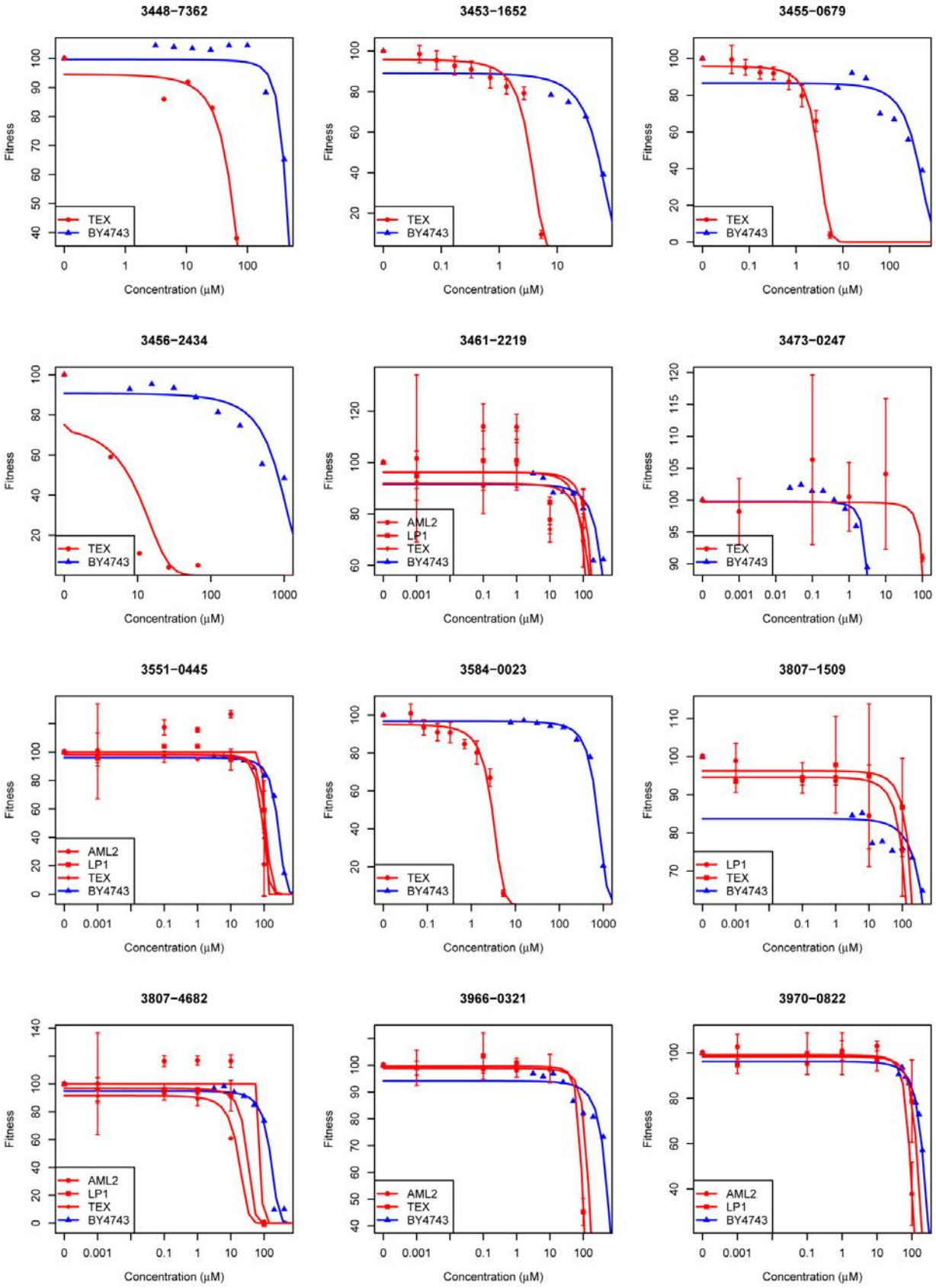

0083−0097
0086−0128
0099−0242
0100−0052
0111−0177
0121−0043
0121−0052
0136−0217
0136−0228
0141−0277
0141−0289
0159−0055
Fitness
Concentration (μM)
BY4743
TEX
AML2
LP1

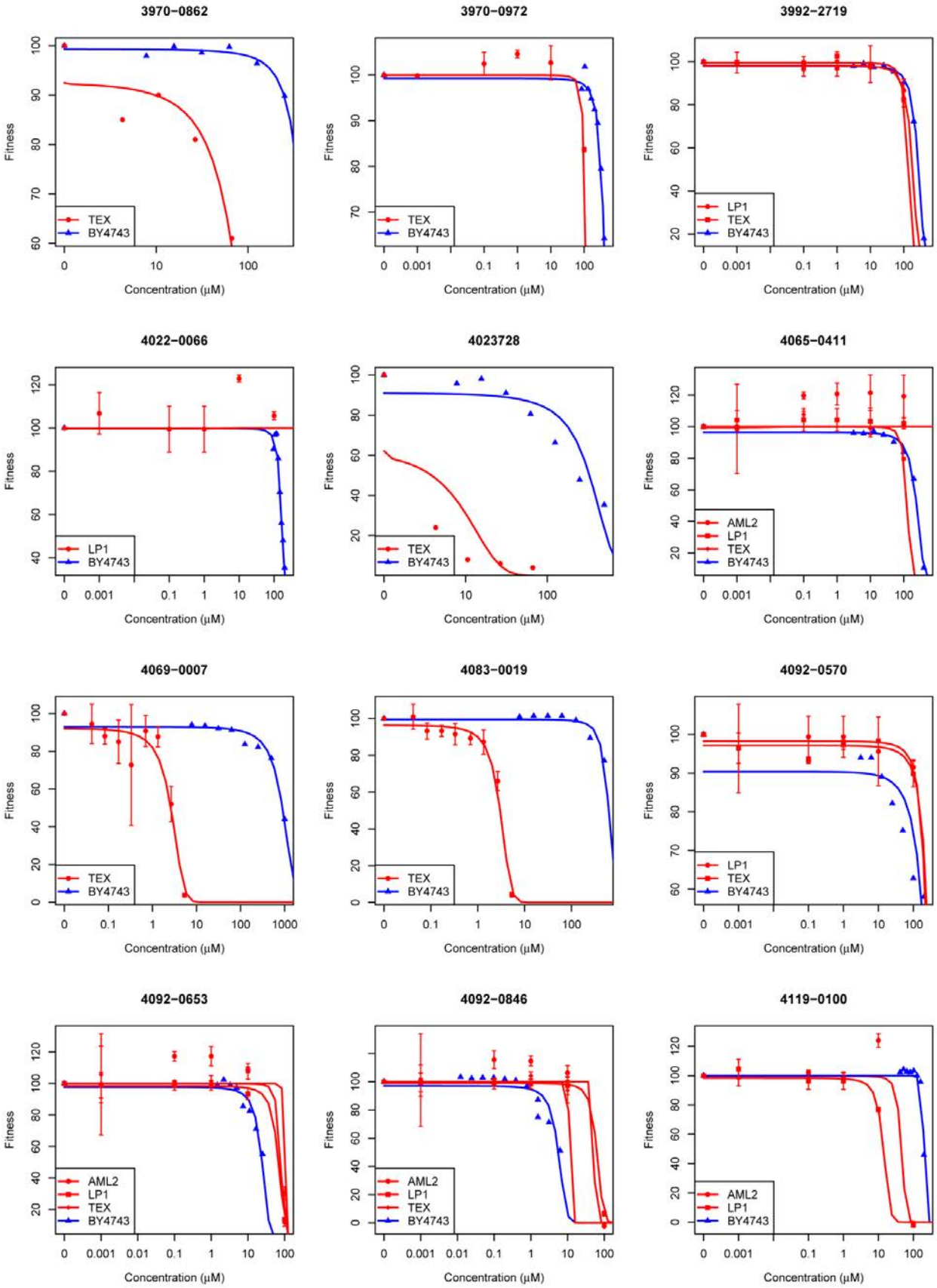

0192−1582
Fitness
Concentration (μM)
TEX
BY4743
0195−0006
AML2
LP1
TEX
BY4743
0269−0018
0279−0214
0330−0160
0337−0584
0418−0143
0438−0332
0469−0687
LP1
0469−0705
0509−0257
0527−0192

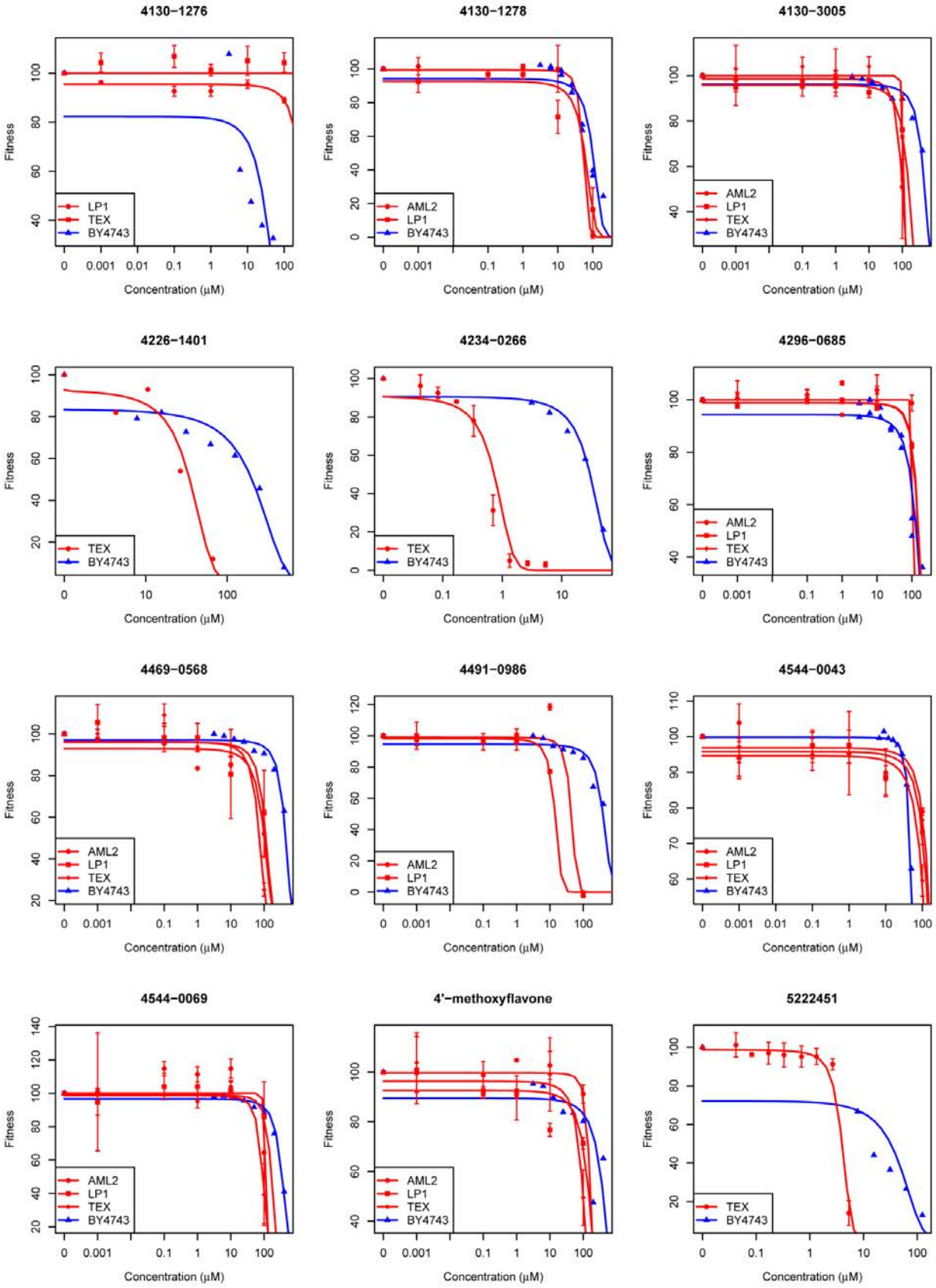

0560−0003
0685−0265
0720−0963
0798−0905
0828−0275
0828−0286
0828−0289
0844−0104
0865−0114
0929−0023
0986−0029
0986−0246
Fitness
Concentration (μM)
BY4743
AML2
LP1
TEX

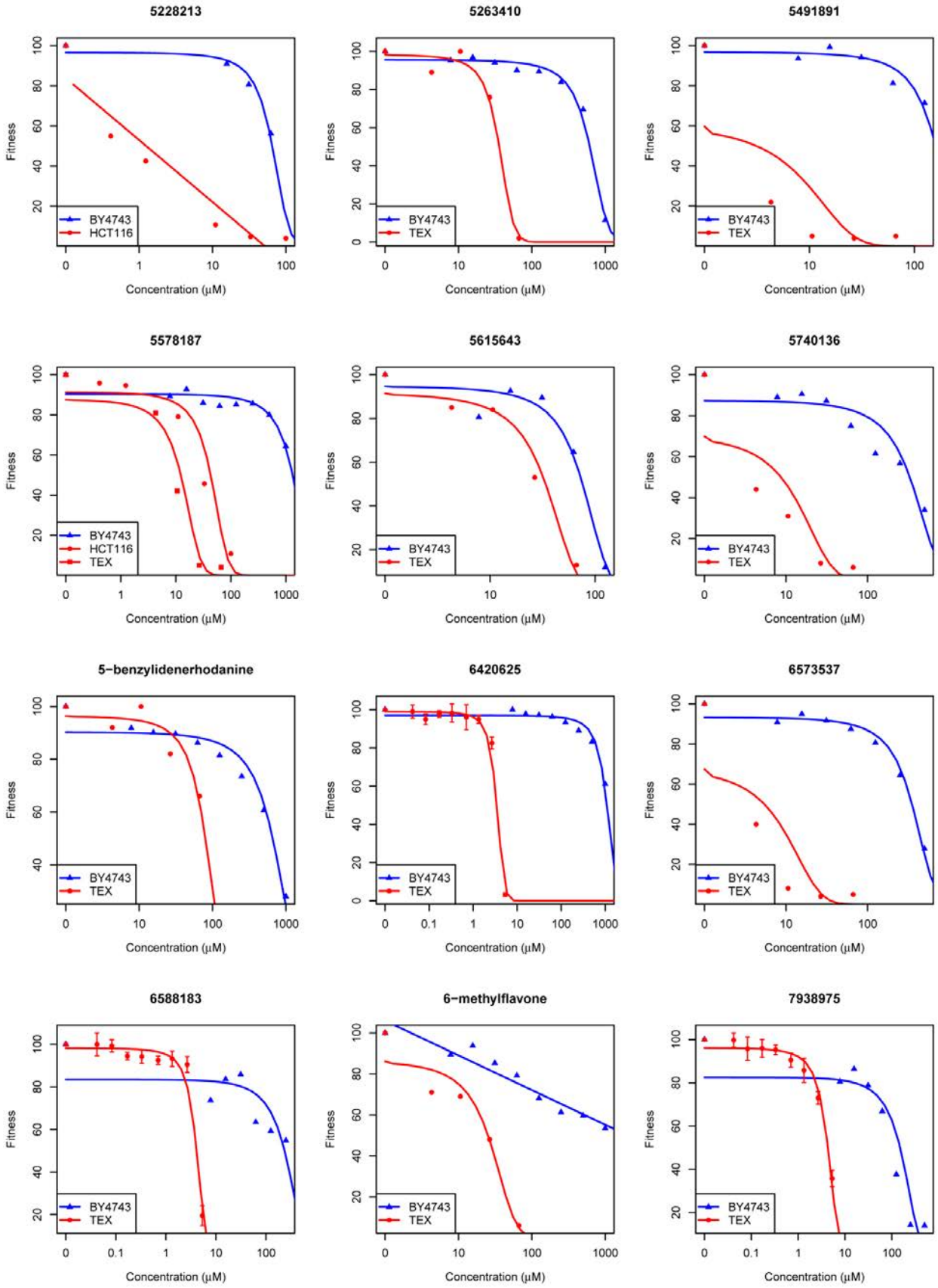

0986−0284
Fitness
Concentration (μM)
TEX
BY4743
0987−0078
AML2
LP1
BY4743
0987−0079
0988−0037
0kpi−0099
1155−0152
1273−0057
1273−0058
1273−0061
1275−0778
1391−0721
1391−0729

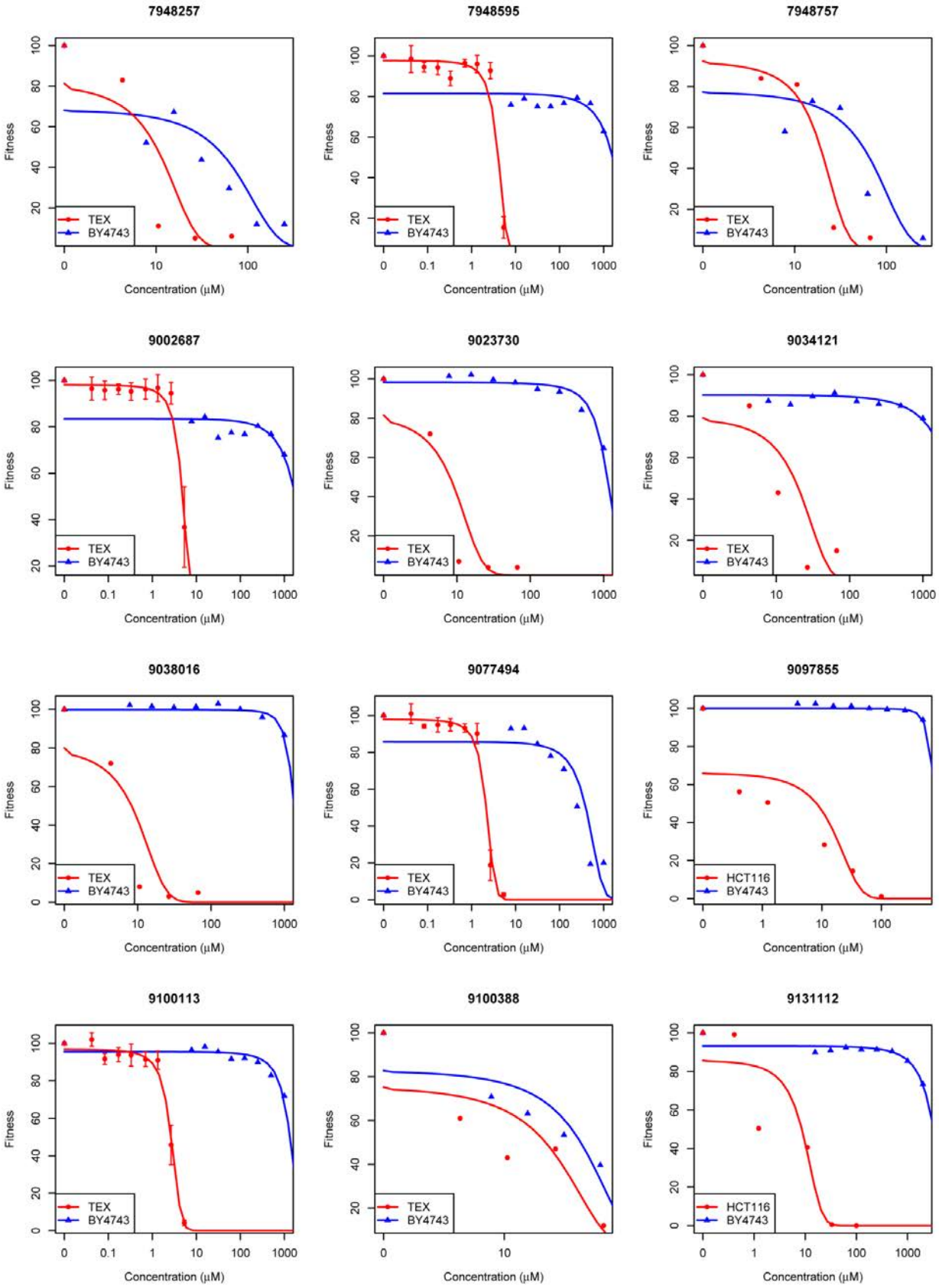

1,3−diallylurea
Fitness
Concentration (μM)
BY4743
HCT116
1436−0052
BY4743
AML2
LP1
TEX
1486−1325
1509−0011
TEX
1543−0184
1548−0686
1594−0068
1671−0008
2324−0395
2421−0008
3013−0144
3346−2049

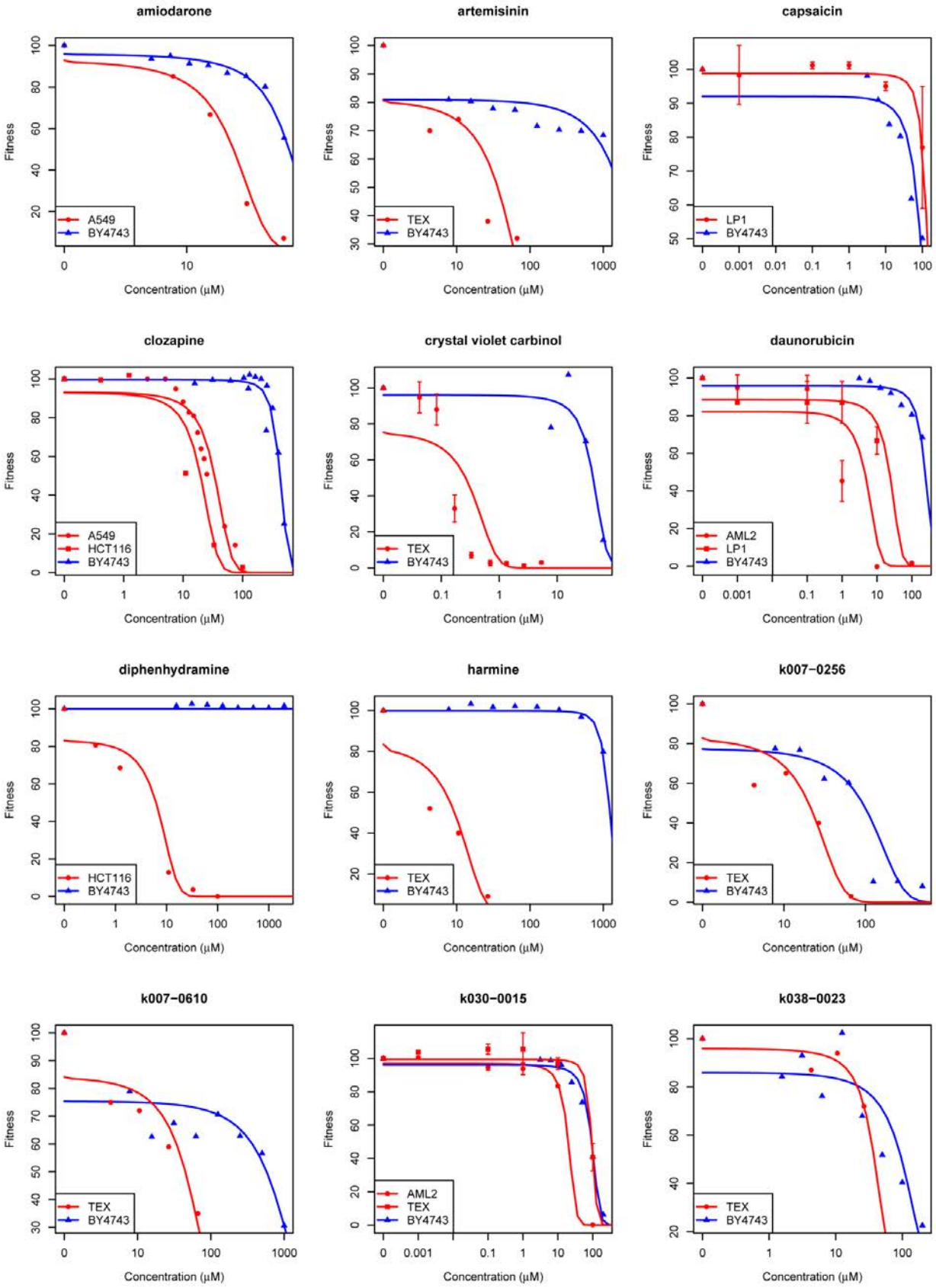

3448−7362
3453−1652
3455−0679
3456−2434
3461−2219
3473−0247
3551−0445
3584−0023
3807−1509
3807−4682
3966−0321
3970−0822
Fitness
Concentration (μM)
TEX
BY4743
AML2
LP1

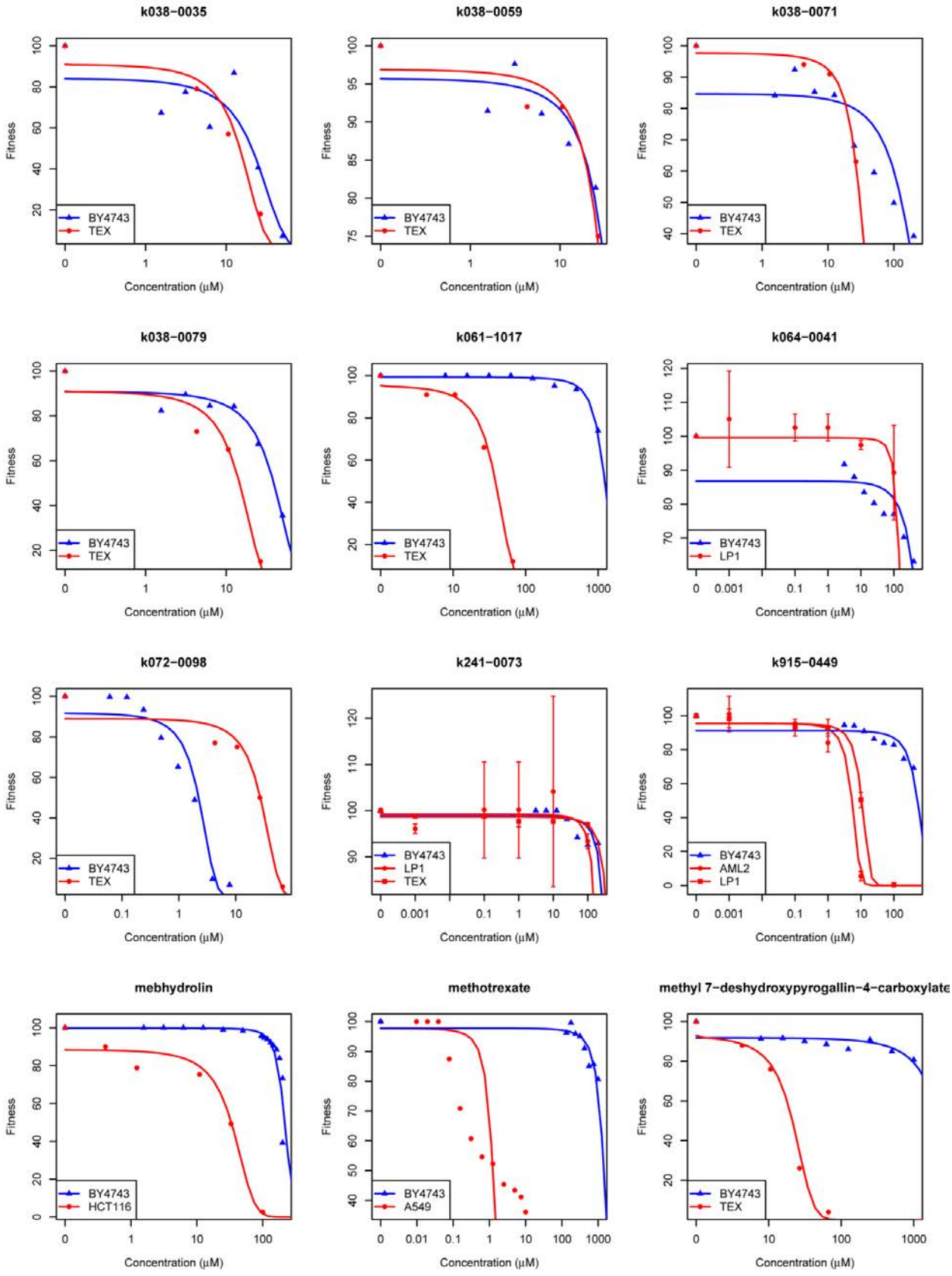

3970−0862
Fitness
Concentration (μM)
TEX
BY4743
3970−0972
Fitness
Concentration (μM)
TEX
BY4743
3992−2719
Fitness
Concentration (μM)
LP1
TEX
BY4743
4022−0066
Fitness
Concentration (μM)
LP1
BY4743
4023728
Fitness
Concentration (μM)
TEX
BY4743
4065−0411
Fitness
Concentration (μM)
AML2
LP1
TEX
BY4743
4069−0007
Fitness
Concentration (μM)
TEX
BY4743
4083−0019
Fitness
Concentration (μM)
TEX
BY4743
4092−0570
Fitness
Concentration (μM)
LP1
TEX
BY4743
4092−0653
Fitness
Concentration (μM)
AML2
LP1
TEX
BY4743
4092−0846
Fitness
Concentration (μM)
AML2
LP1
TEX
BY4743
4119−0100
Fitness
Concentration (μM)
AML2
LP1
BY4743


4130−1276
Fitness
Concentration (μM)
LP1
TEX
BY4743
4130−1278
AML2
LP1
BY4743
4130−3005
AML2
LP1
TEX
BY4743
4226−1401
TEX
BY4743
4234−0266
TEX
BY4743
4296−0685
AML2
LP1
TEX
BY4743
4469−0568
AML2
LP1
TEX
BY4743
4491−0986
AML2
LP1
BY4743
4544−0043
AML2
LP1
TEX
BY4743
4544−0069
AML2
LP1
TEX
BY4743
4'−methoxyflavone
AML2
LP1
TEX
BY4743
5222451
TEX
BY4743


5228213
Fitness
Concentration (μM)
BY4743
HCT116
5263410
BY4743
TEX
5491891
BY4743
TEX
5578187
BY4743
HCT116
TEX
5615643
BY4743
TEX
5740136
BY4743
TEX
5−benzylidenerhodanine
BY4743
TEX
6420625
BY4743
TEX
6573537
BY4743
TEX
6588183
BY4743
TEX
6−methylflavone
BY4743
TEX
7938975
BY4743
TEX


7948257
7948595
7948757
9002687
9023730
9034121
9038016
9077494
9097855
9100113
9100388
9131112
Fitness
Concentration (µM)
TEX
BY4743
HCT116


amiodarone
Fitness
Concentration (μM)
A549
BY4743
artemisinin
TEX
BY4743
capsaicin
LP1
BY4743
clozapine
A549
HCT116
BY4743
crystal violet carbinol
TEX
BY4743
daunorubicin
AML2
LP1
BY4743
diphenhydramine
HCT116
BY4743
harmine
TEX
BY4743
k007−0256
TEX
BY4743
k007−0610
TEX
BY4743
k030−0015
AML2
TEX
BY4743
k038−0023
TEX
BY4743


k038−0035
k038−0059
k038−0071
k038−0079
k061−1017
k064−0041
k072−0098
k241−0073
k915−0449
mebhydrolin
methotrexate
methyl 7−deshydroxypyrogallin−4−carboxylate
Fitness
Concentration (μM)
BY4743
TEX
LP1
AML2
HCT116
A549

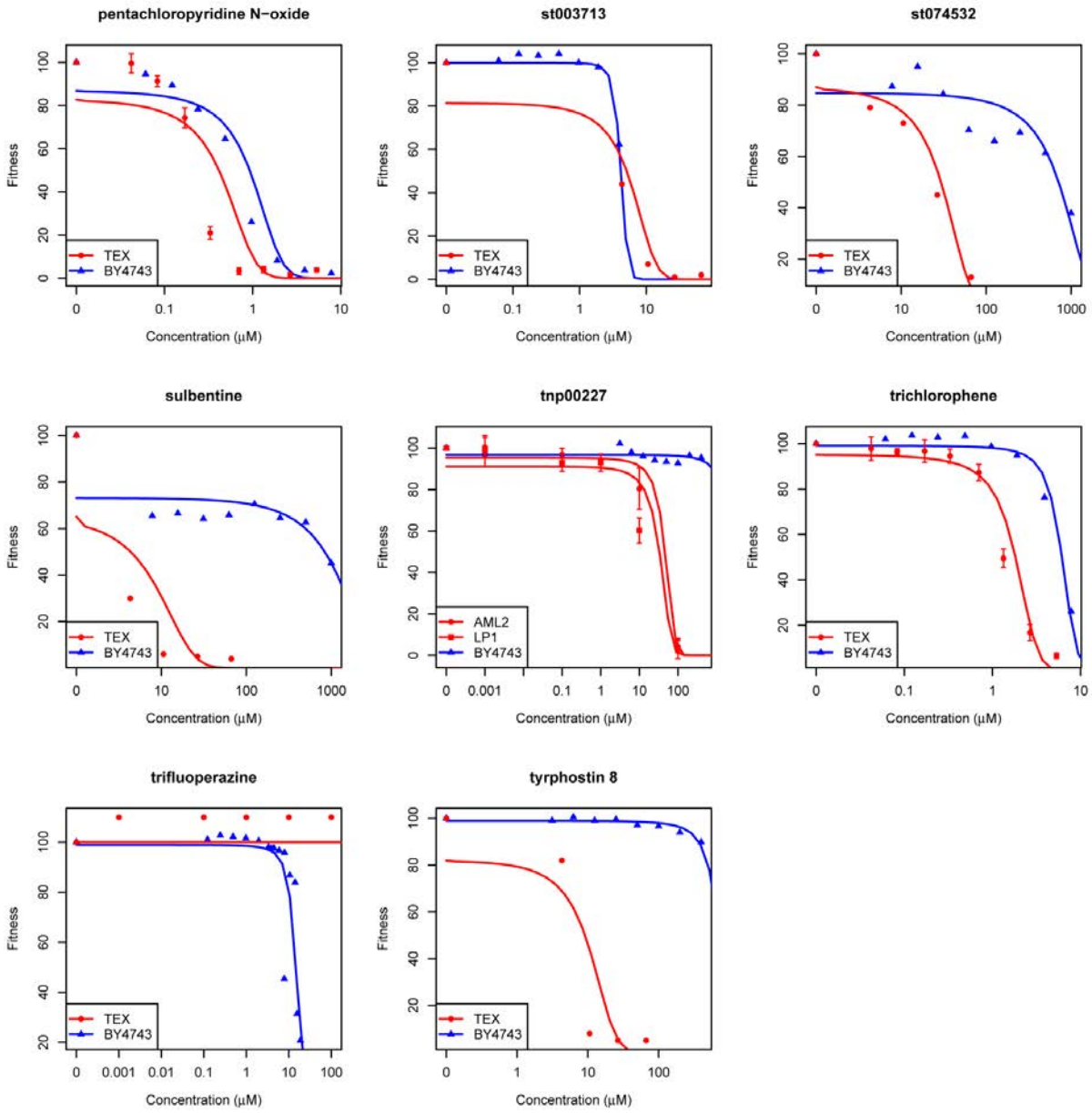


Fig. S4

Dose-response curves of yeast (BY4743) versus mammalian cells. Each plot shows the dose-response curves of yeast (blue) and mammalian cells (red) treated with one compound (indicated above the plot), where the response is growth fitness expressed as a percentage of the fitness of DMSO-treated cells (i.e. mock-treated cells). Growth of the wildtype BY4743 *hoΔ/HO* control strain was monitored (OD_{600}) every 15 minutes, allowing estimates of the doubling time under different conditions. Mammalian viability assays are described in (*10*). Fitness of the yeast cells was estimated as <doubling time of DMSO-treated cells> / <doubling time of compound-treated cells> x 100%. Fitness of mammalian cells was estimated as <# of compound-treated cells> / <# of DMSO-treated cells> x100%. Where replicates are available, data are mean +/- SD ($N \geq 3$). For each curve, a fitted line is also shown (sigmoid in most cases, linear in cases where sigmoid fits are poor). Dose-response curves are shown for a total of 152 compounds.

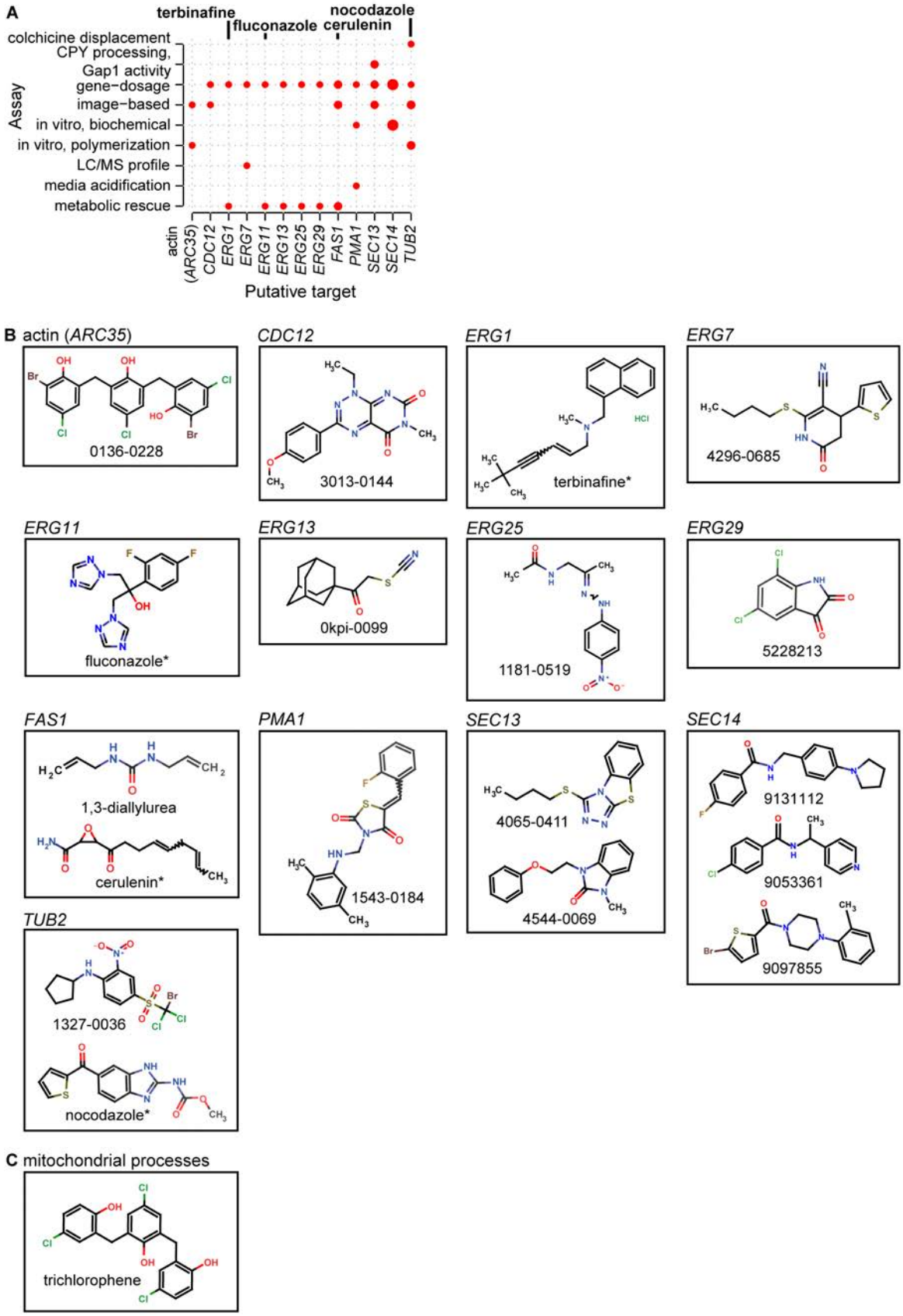
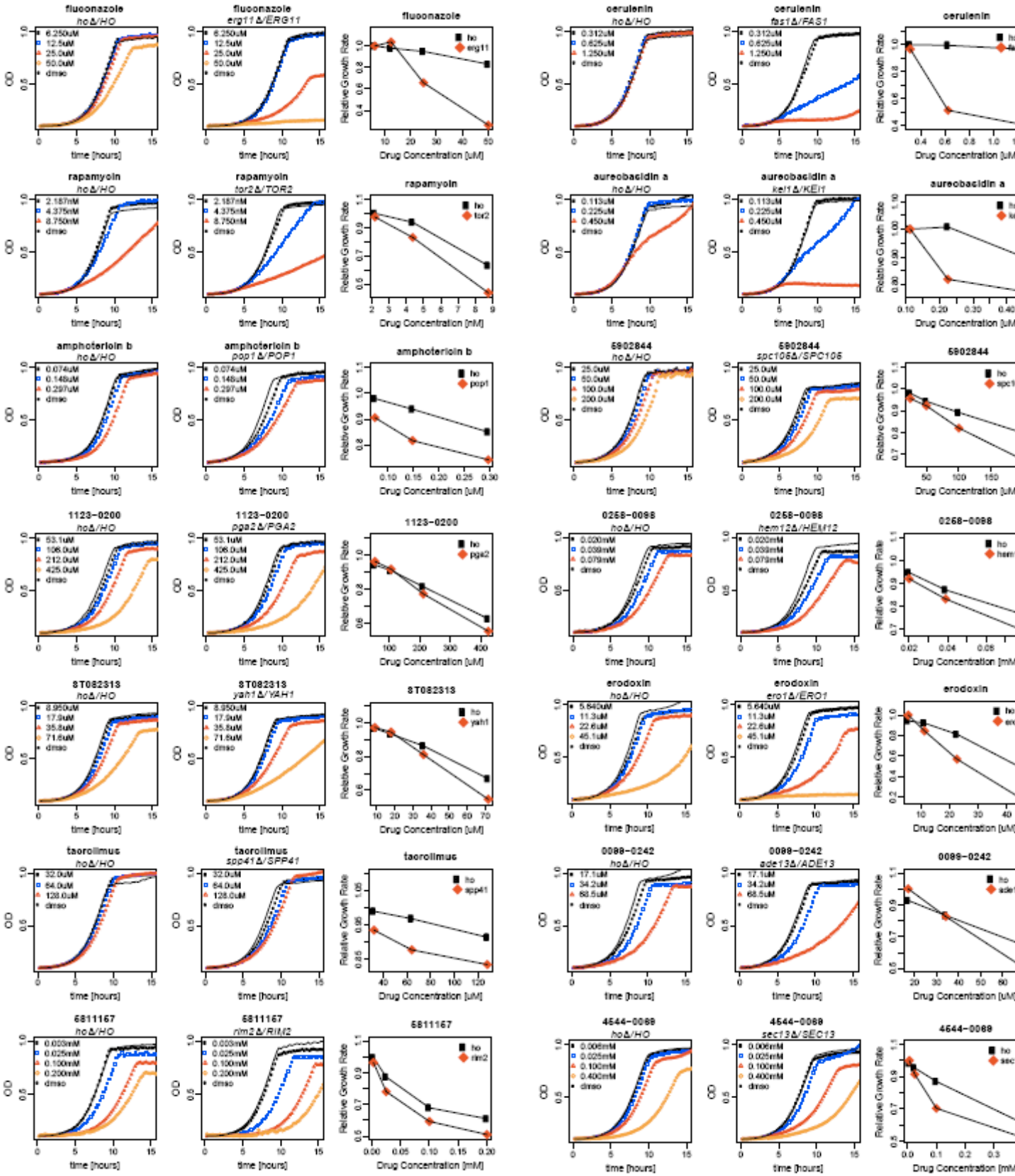
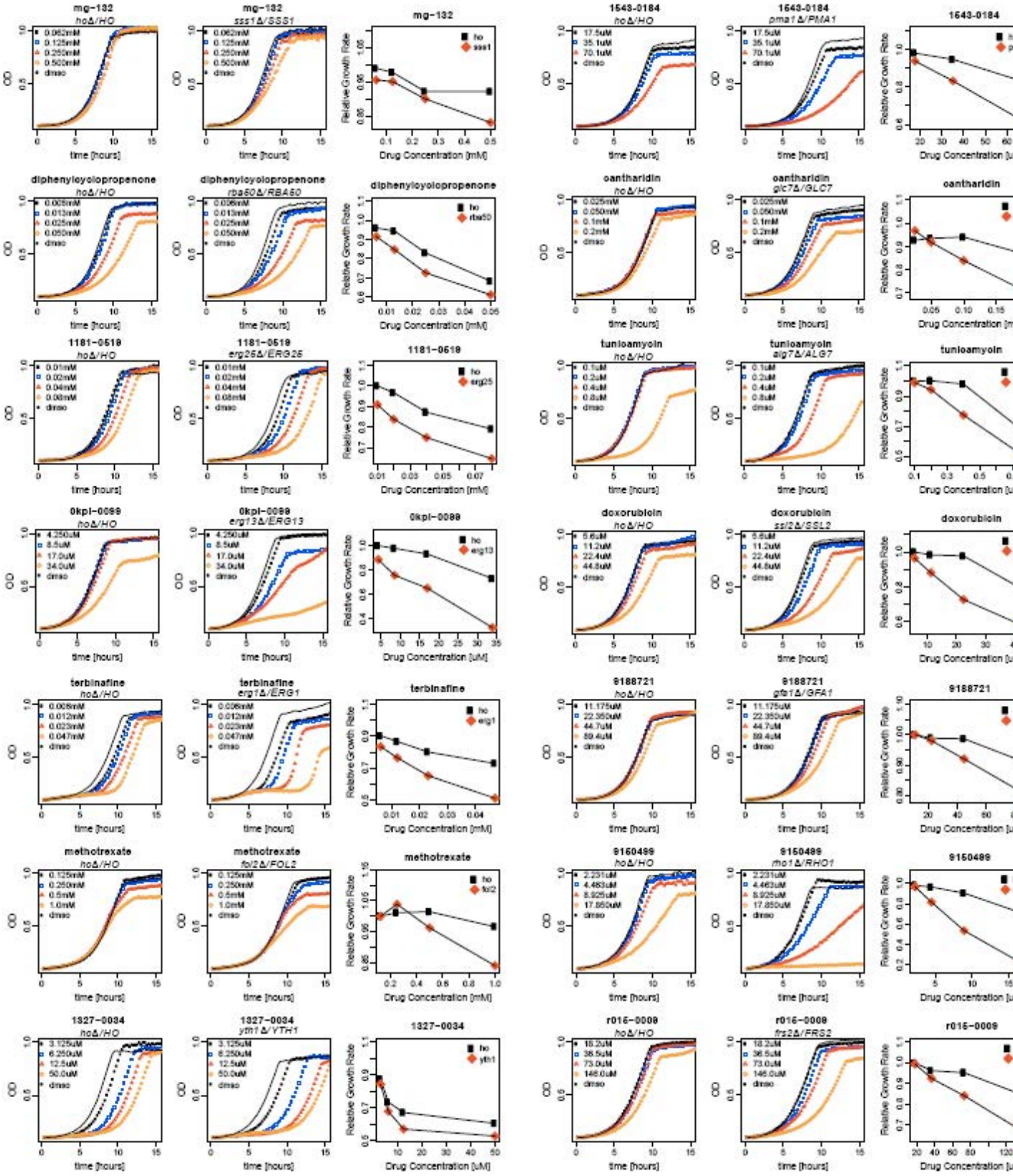
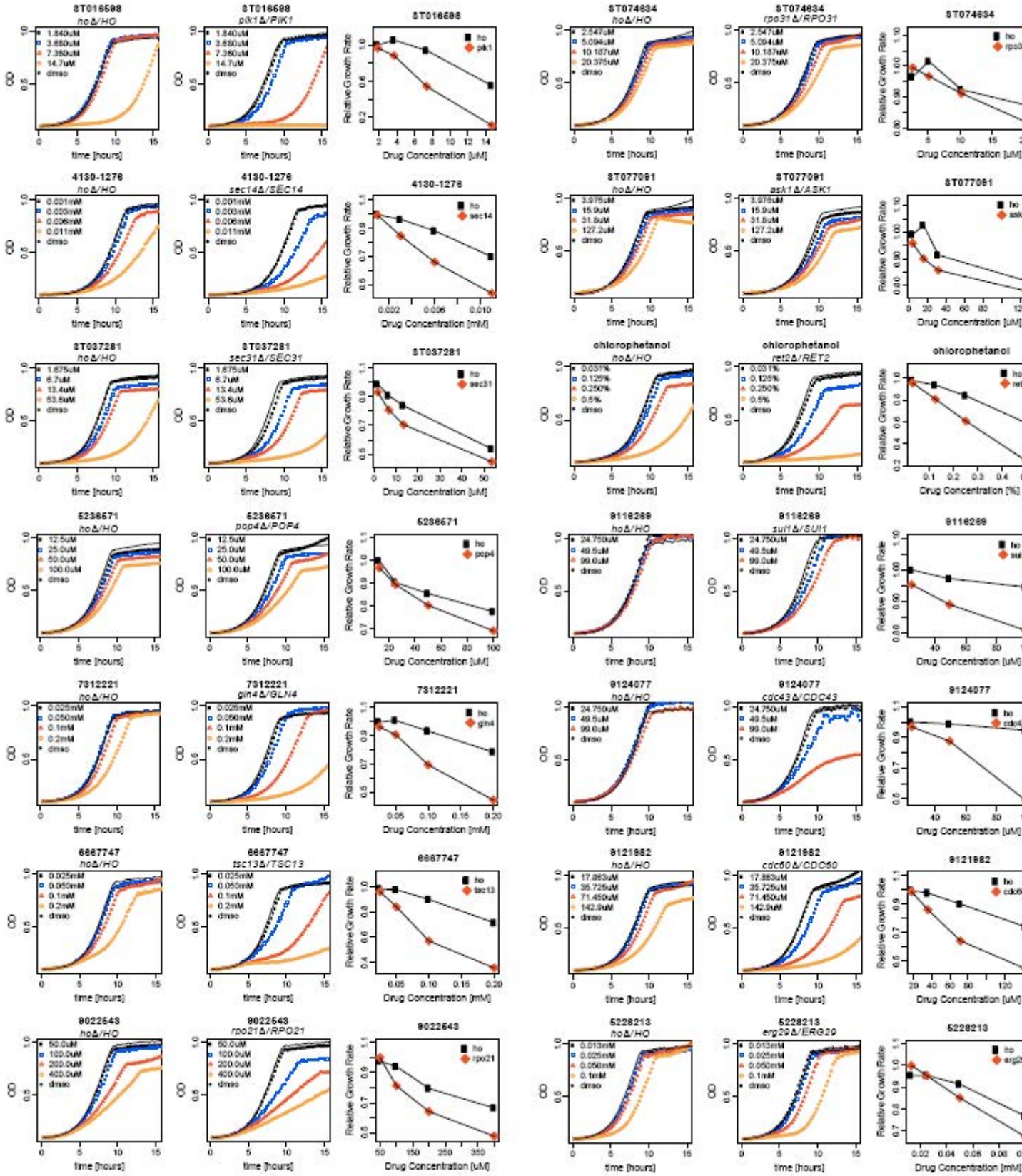

A
terbinafine
fluconazole
nocodazole
cerulenin
colchicine displacement
CPY processing,
Gap1 activity
gene−dosage
image−based
in vitro, biochemical
in vitro, polymerization
LC/MS profile
media acidification
metabolic rescue
Assay
actin (ARC35)
CDC12
ERG1
ERG7
ERG11
ERG13
ERG25
ERG29
FAS1
PMA1
SEC13
SEC14
TUB2
Putative target
B actin (ARC35)
0136-0228
CDC12
3013-0144
ERG1
terbinafine*
ERG7
4296-0685
ERG11
fluconazole*
ERG13
0kpi-0099
ERG25
1181-0519
ERG29
5228213
FAS1
1,3-diallylurea
cerulenin*
PMA1
1543-0184
SEC13
4065-0411
4544-0069
SEC14
9131112
9053361
9097855
TUB2
1327-0036
nocodazole*
C mitochondrial processes
trichlorophene

Fig. S5 (previous page)

Summary of validated inhibitors. (A) Validation assays of chemical-genetic probes. Points indicate assays performed (y-axis) for putative compound targets (x-axis); size proportional to number of compounds validated. Compounds with established targets are indicated above the plot. Validation assays are provided for actin and *TUB2* (**fig. S10**), *CDC12* (**Fig. 1B**), *ERG* genes (**fig. S12**), *FAS1* (**fig. S9**), *PMA1* (**fig. S23**), *SEC13* (**fig. S24**) and *SEC14* (**Figs. 1C, S7**). **(B)** Chemical structures of the chemical-genetic probes that were validated as protein inhibitors, with the assays indicated in panel A. Inhibitors with the same putative target are shown together in the same box, with the target indicated above the box. *, compounds with established targets. **(C)** Chemical structure of trichlorophene, the putative inhibitor of mitochondrial processes, predicted by its response signature, and validated as shown in **fig. S13**.


fluconazole
ho Δ/HO
fluconazole
erg11Δ/ERG11
fluconazole
ho
erg11
Relative Growth Rate
Drug Concentration [uM]
6.250uM
12.5uM
25.0uM
50.0uM
dmso
time [hours]
OD
cerulenin
ho Δ/HO
cerulenin
fas1Δ/FAS1
cerulenin
0.312uM
0.625uM
1.250uM
rapamycin
ho Δ/HO
rapamycin
tor2Δ/TOR2
rapamycin
tor2
2.187nM
4.375nM
8.750nM
Drug Concentration [nM]
aureobasidin a
ho Δ/HO
aureobasidin a
kei1Δ/KEI1
aureobasidin a
0.113uM
0.225uM
0.450uM
amphotericin b
ho Δ/HO
amphotericin b
pop1Δ/POP1
amphotericin b
pop1
0.074uM
0.148uM
0.297uM
5902844
ho Δ/HO
5902844
spc105Δ/SPC105
5902844
25.0uM
50.0uM
100.0uM
200.0uM
1123−0200
ho Δ/HO
1123−0200
pga2Δ/PGA2
1123−0200
pga2
53.1uM
106.0uM
212.0uM
425.0uM
0258−0088
ho Δ/HO
0258−0088
hem12Δ/HEM12
0258−0088
0.020mM
0.039mM
0.079mM
8T082313
ho Δ/HO
8T082313
yah1Δ/YAH1
8T082313
yah1
8.950uM
17.9uM
35.8uM
71.6uM
erodoxin
ho Δ/HO
erodoxin
ero1Δ/ERO1
erodoxin
5.640uM
11.3uM
22.6uM
45.1uM
taorolimus
ho Δ/HO
taorolimus
spp41Δ/SPP41
taorolimus
spp41
32.0uM
64.0uM
128.0uM
0088−0242
ho Δ/HO
0088−0242
ade13Δ/ADE13
0088−0242
17.1uM
34.2uM
68.5uM
6811167
ho Δ/HO
6811167
rim2Δ/RIM2
6811167
rim2
0.025mM
0.100mM
0.200mM
4544−0069
ho Δ/HO
4544−0069
sec13Δ/SEC13
4544−0069
0.006mM
0.400mM


mg-132
ho∆/HO
mg-132
sss1∆/SSS1
mg-132
1643-0184
ho∆/HO
1643-0184
pma1∆/PMA1
1643-0184
diphenyleyelopropenone
ho∆/HO
diphenyleyelopropenone
rba60∆/RBA60
diphenyleyelopropenone
oantharidin
ho∆/HO
oantharidin
glc7∆/GLC7
oantharidin
1181-0518
ho∆/HO
1181-0518
erg25∆/ERG25
1181-0518
tunicamycin
ho∆/HO
tunicamycin
alg7∆/ALG7
tunicamycin
0kpi-0088
ho∆/HO
0kpi-0088
erg13∆/ERG13
0kpi-0088
doxorubicin
ho∆/HO
doxorubicin
ssl2∆/SSL2
doxorubicin
terbinafine
ho∆/HO
terbinafine
erg1∆/ERG1
terbinafine
9188721
ho∆/HO
9188721
gfa1∆/GFA1
9188721
methotrexate
ho∆/HO
methotrexate
fol2∆/FOL2
methotrexate
9150499
ho∆/HO
9150499
rho1∆/RHO1
9150499
1327-0034
ho∆/HO
1327-0034
yth1∆/YTH1
1327-0034
r016-0009
ho∆/HO
r016-0009
frs2∆/FRS2
r016-0009
time [hours]
Drug Concentration [mM]
Drug Concentration [uM]
Relative Growth Rate
OD
ho
sss1
rba60
erg25
erg13
erg1
fol2
yth1
pma1
glc7
alg7
ssl2
gfa1
rho1
frs2


ST016698
hoΔ/HO
pik1Δ/PIK1
ST074834
rpo31Δ/RPO31
4130-1278
sec14Δ/SEC14
ST077081
ask1Δ/ASK1
ST037281
sec31Δ/SEC31
ohlorophetanol
ret2Δ/RET2
5236671
pop4Δ/POP4
8116289
sui1Δ/SUI1
7312221
gln4Δ/GLN4
8124077
cdc43Δ/CDC43
8667747
tsc13Δ/TSC13
8121982
cdc60Δ/CDC60
8022643
rpo21Δ/RPO21
5228213
erg29Δ/ERG29
time [hours]
Drug Concentration [uM]
Drug Concentration [mM]
Drug Concentration [%]
Relative Growth Rate
OD
dmso

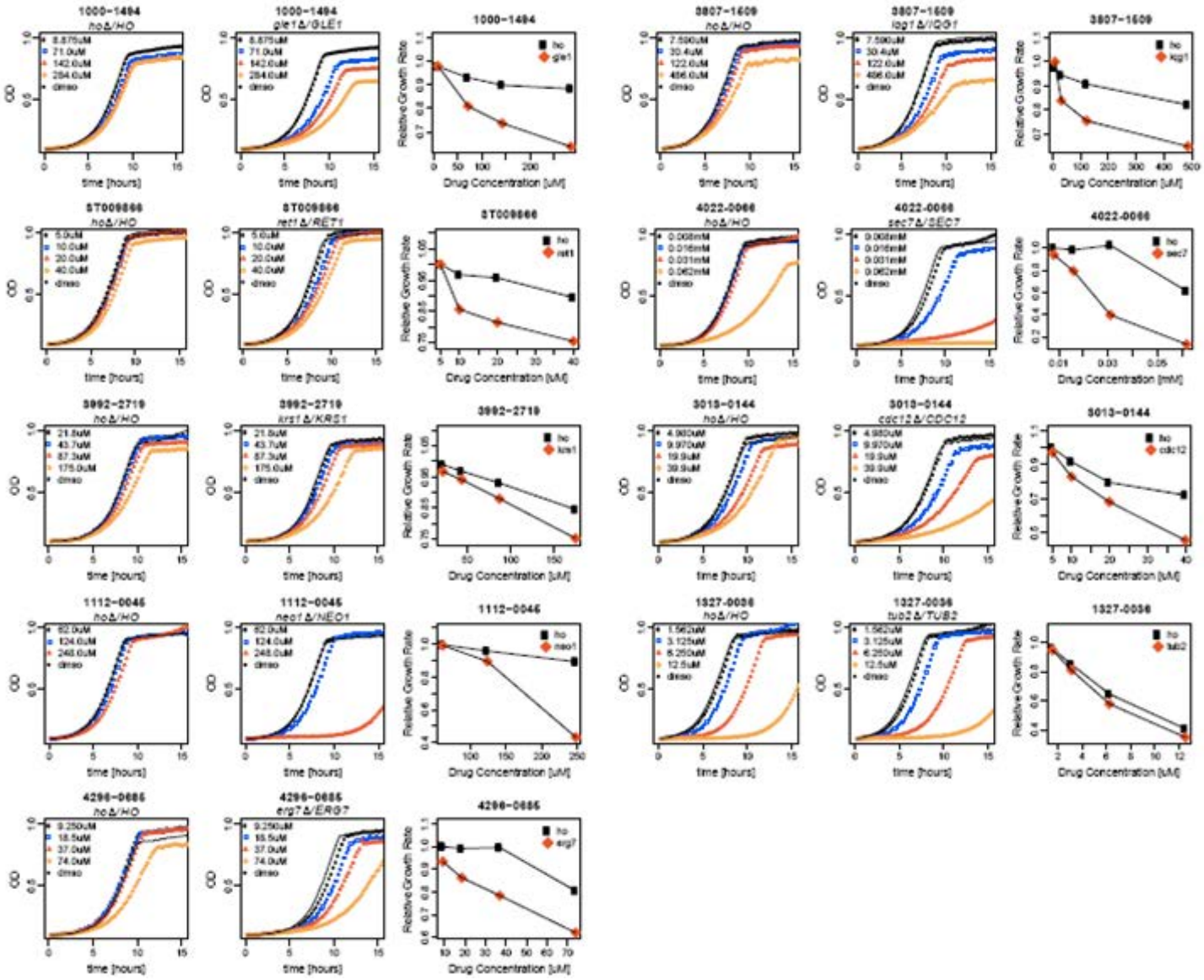


Fig. S6

Validation of drug-induced haploinsuffiency for 51 compound-strain pairs by individual strain analysis. Significantly scoring heterozygous deletion strains, 44 of which were designated as HIP hits, in individual screens were validated by confirming the expected drug-induced haploinsufficiency phenotype of heterozygous deletion strains grown individually in the presence of compound. Results for each compound-strain pair are presented in three plots. The growth of the heterozygous deletion strain (OD_{600}) (middle plot, strain name indicated above) and that of the *hoΔ/HO* control strain (left plot) were monitored in various concentrations of the compound (see legend). Dose-response curves summarizing these data are shown in the right plots. Relative growth rates (y-axis) of the heterozygous strain (red) and the control strain (black) at increasing concentrations of compound (x-axis) are shown. Relative growth rate was calculated as the average doubling time in DMSO (vehicle) divided by that of the compound-treated culture. Validation of drug-induced haploinsufficiency for an additional three *SEC14*, eight *NEO1* and one *FAS1* compound-strain pairs are included as parts of **figs. S7, S8** and **S9**, respectively.

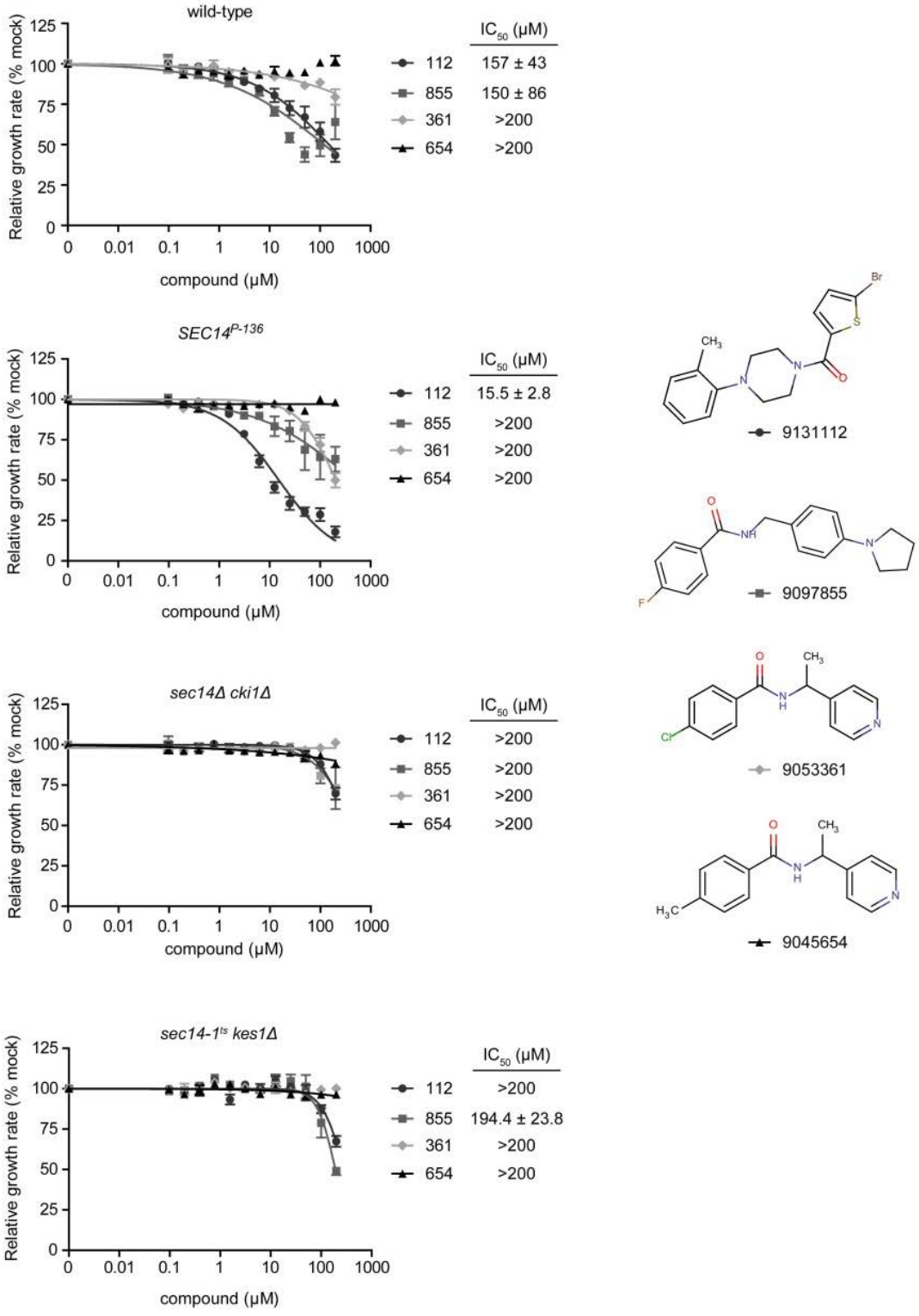


Fig. S7

Genetic validation of Sec14 inhibitors. Results shown demonstrate increased sensitivity to a Sec14 knockdown mutant (*SEC14*$^{P\text{-}136}$) and increased resistance to Sec14 bypass mutants (*cki1Δ* and *kes1Δ*). Growth rate relative to the no-compound control (y-axis) was measured at increasing concentrations of 9131112, 9097855, 9053361 and 9045654 (an inactive analog) (x-axis).

Results for wild-type (CTY182; top panel), a *SEC14* promoter mutant (CTY374; $SEC14^{P\text{-}136}$) in which steady-state Sec14 protein levels are reduced ~7-fold relative to wild-type (second panel), the *SEC14* bypass mutants, *cki1Δ* (CTY303; third panel) and *kes1Δ* (CTY159; bottom panel) are shown. Data are mean +/- SEM ($N = 3$). IC_{50}s, i.e. 50% inhibitory concentrations, represent the 95% confidence interval. Chemical structures are shown on the right.

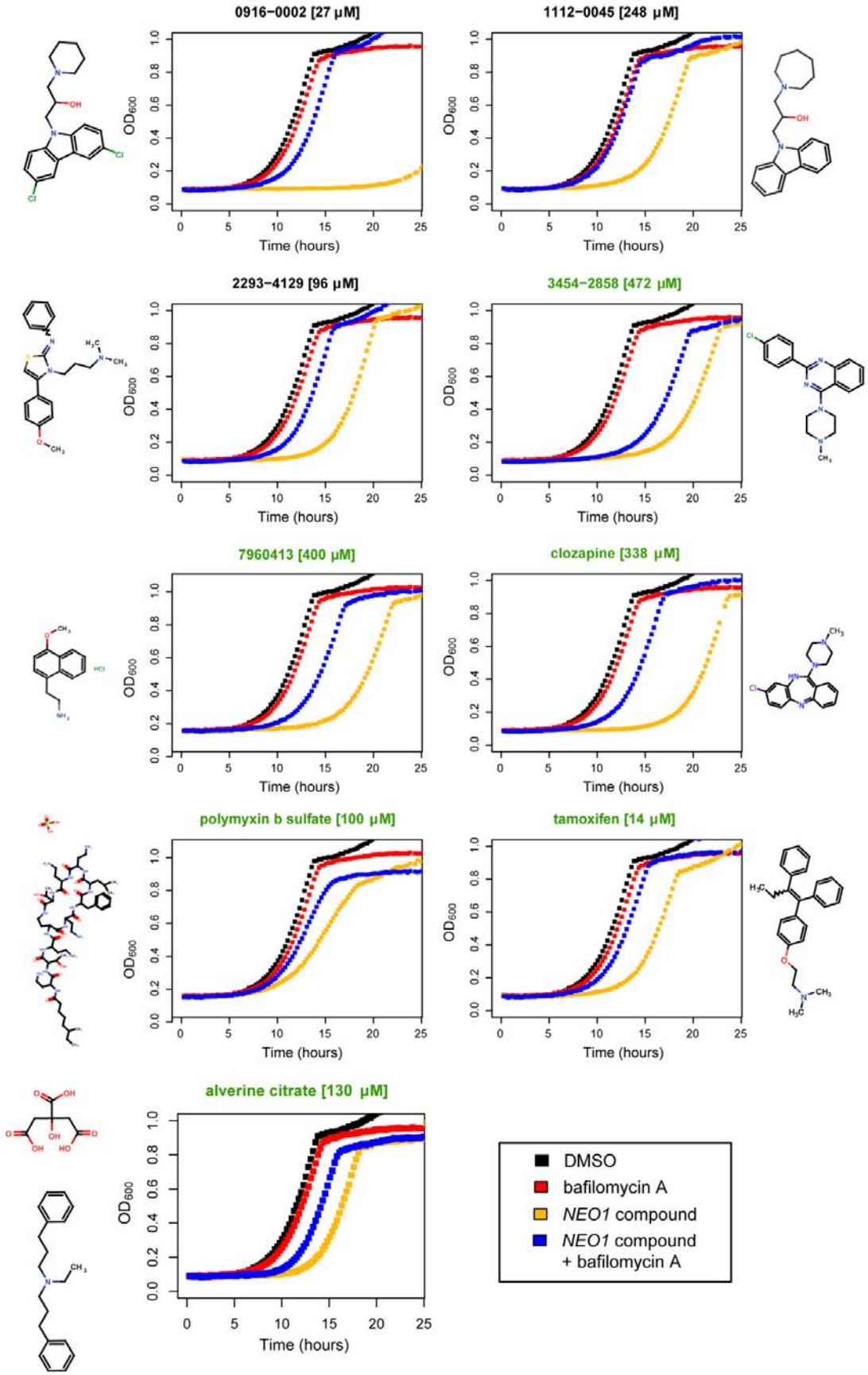

0916−0002 [27 µM]
1112−0045 [248 µM]
2293−4129 [96 µM]
3454−2858 [472 µM]
7960413 [400 µM]
clozapine [338 µM]
polymyxin b sulfate [100 µM]
tamoxifen [14 µM]
alverine citrate [130 µM]
OD600
Time (hours)
DMSO
bafilomycin A
NEO1 compound
NEO1 compound + bafilomycin A

Fig. S8 (previous page)

Bafilomycin A rescues the drug-induced NEO1 haploinsufficiency of compounds with NEO1 as a HIP hit. Growth of the *neo1*Δ*/NEO1* strain was monitored by measuring the OD_{600} (y-axis) every 15 minutes for 24h (x-axis). In each plot, the growth of *neo1*Δ*/NEO1* under four conditions is presented: DMSO (black), 1.25μM bafilomycin A (red), a *NEO1* compound (yellow, the name of the compound is above the plot), and a *NEO1* compound + 1.25μM bafilomycin A (blue). Each *NEO1* compound induces the *NEO1* or *NEO1-PIK1* signature except alverine citrate, compounds that are also cationic amphiphilic drugs (CADs) are indicated in green, and the concentration used in the experiment is shown in square brackets above the corresponding plot. The molecular structure of each *NEO1* compound is illustrated beside each plot.

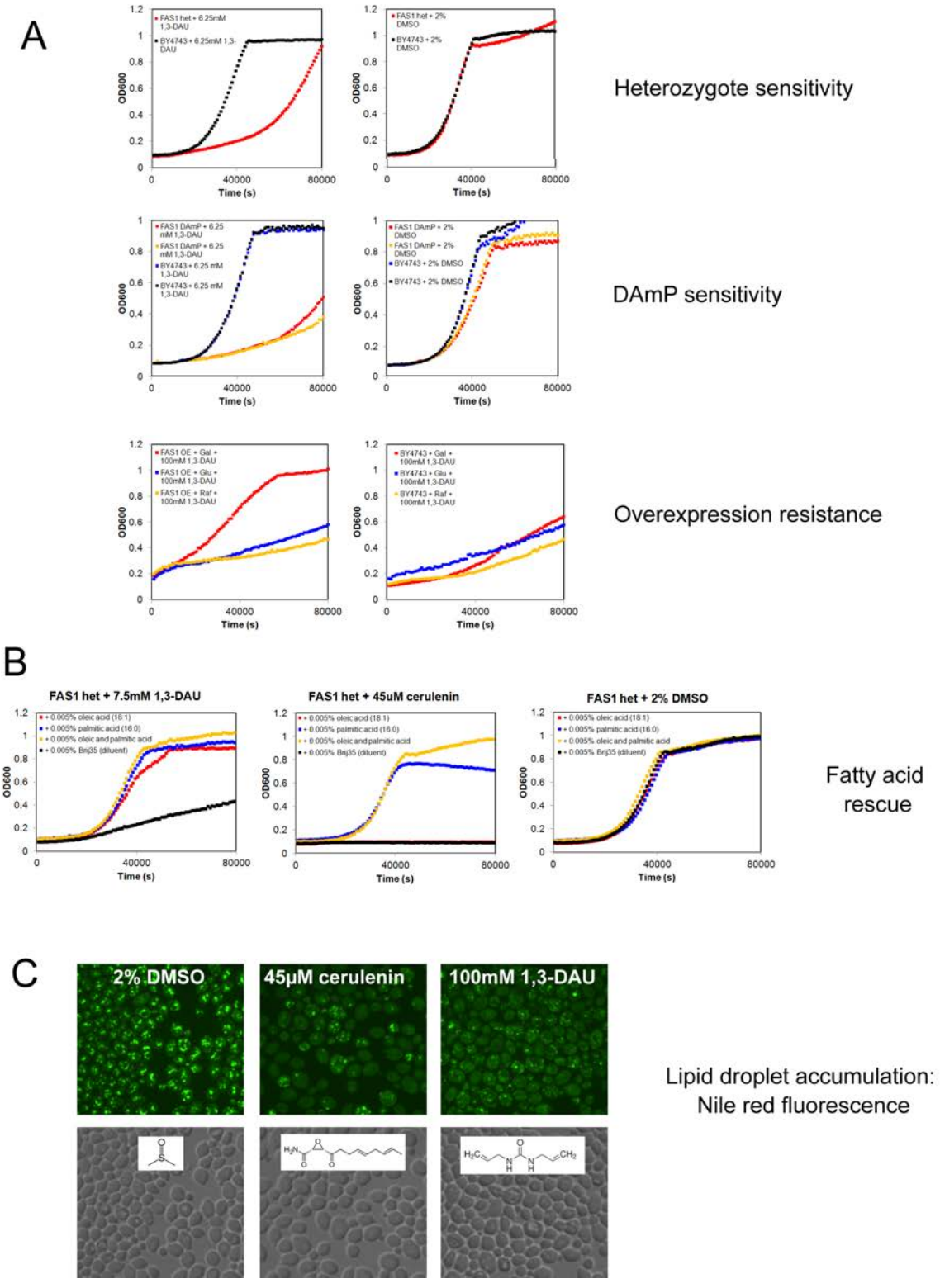


Fig. S9

Validation of Fas1 inhibitors cerulenin and 1,3 diallylurea. (A) Growth curves of *FAS1* mutant strains in the presence of 1,3-diallylurea (1,3-DAU). *FAS1* heteroyzygous, *FAS1* DAmP, and BY4743 background strains were grown in YPD media, in the presence of 6.25mM 1,3-DAU or 2% DMSO (vehicle). The *FAS1* overexpression strain and BY4743 background strain

were grown in synthetic media containing galactose (Gal; to induce overexpression), glucose (Glu; to repress overexpression), or raffinose (Raf; neutral carbon source) in the presence of 100mM 1,3-DAU. **(B)** Fatty acids rescue. The *FAS1* heterozygote strain was grown in the presence of 7.5mM 1,3-DAU and 177μM oleic acid, or 195μM palmitic acid, or a 50:50 mixture of oleic and palmitic acid (186 μM), or 0.005% (w/v) Brij-35 (diluent). The fatty acid rescue experiment was also performed in the presence of the fatty acid biosynthesis inhibitor cerulenin (45μM) or 2% DMSO (vehicle). **(C)** 1,3-DAU leads to decreased lipid droplet formation. Wild-type BY4743 cells were grown in the presence of 100mM 1,3-DAU, 45μM cerulenin, or 2% DMSO for 6h. Cells were stained with Nile Red to visualize lipid droplets. Briefly cells were resuspended in IM sorbitol, fixed in 1% formaldehyde for 5min, washed in 50mM Tris-HCl ph7.5 and stained with 2ug/mL Nile Red for 10min before imaging.

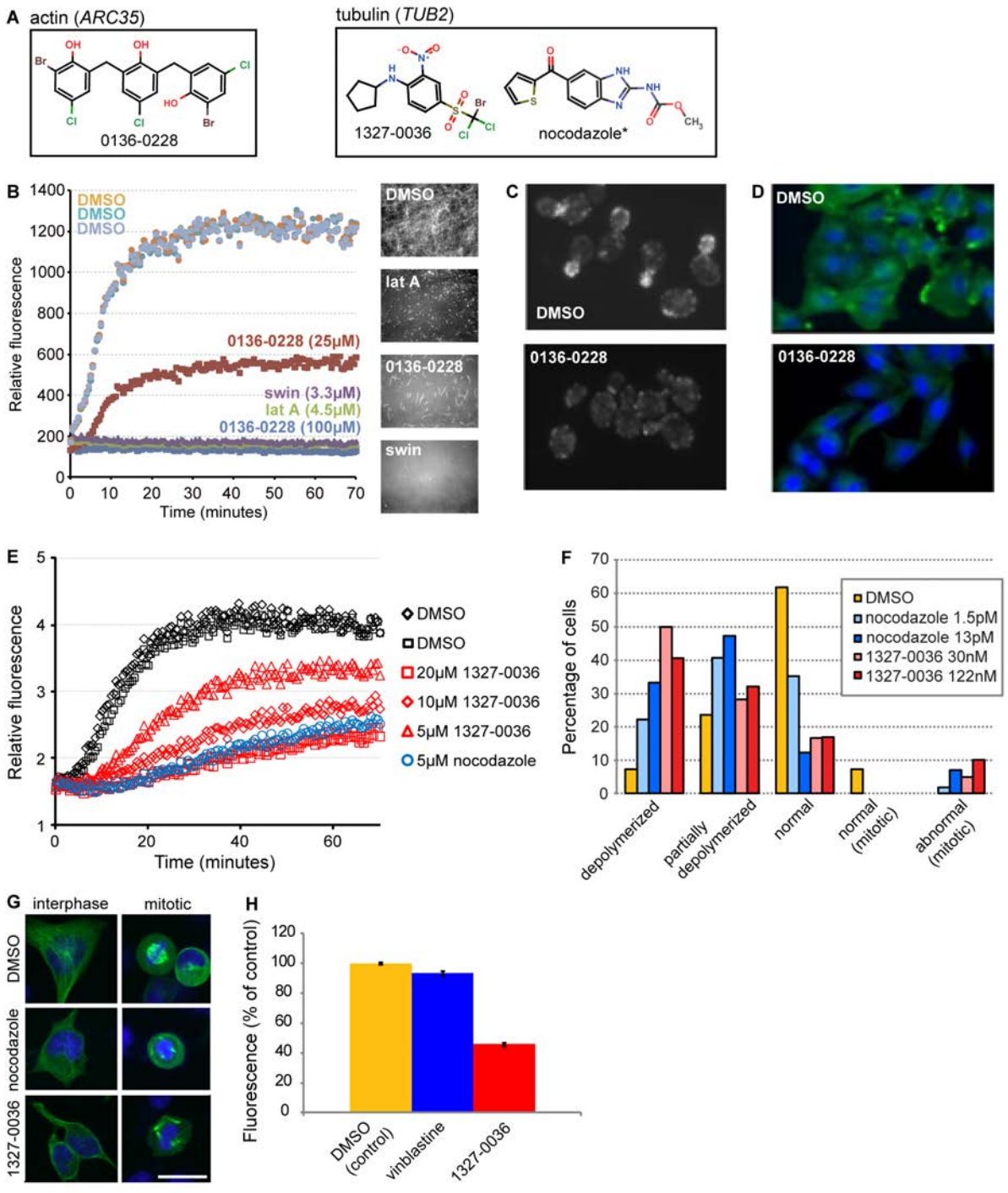


Fig. S10

Confirmation assays of actin and tubulin inhibitors. (A) Chemical structures of inhibitors that were HIPHOP profiled and tested for their ability to inhibit actin/tubulin. *, *TUB2* is one of established targets of nocodazole. **(B)** *In vitro* confirmation assay of the 0136-0228 actin inhibitor at 25μM and 100μM. In this assay, actin monomer is conjugated to pyrene and the monomeric pyrene G actin allowed to polymerize. The resulting pyrene f-actin manifests a greatly increased fluorescent yield. The course of polymerization is monitored in a fluorescent spectrophotometer to compare the relative fluorescent yield over time in the presence of 0136-

0228, DMSO as a negative control, or with swinholide and latrunculin A as positive controls. These exact same polymerization reactions were observed microscopically and are shown as a series of four images. In DMSO, a dense network of actin filaments is observed, while in latrunculin A (lat A) very short actin seeds are observed. In 0136-0228, short, aborted actin filament fragments are observed, consistent with a partial inhibition of polymerization. In swinholide (swin), no evidence of filaments is observed. **(C)** Yeast F-actin images show wildtype haploid yeast cells treated with DMSO (top) or 100μM of the actin inhibitor (bottom) (0136-0228) stained with fluorescent phalloidin. DMSO-treated cells show bright actin patches at bud tips and septa whereas actin inhibitor-treated cells lack actin foci and show diffuse cytoplasmic staining. **(D)** Mammalian cells (A549) stained with phalloidin show a robust network of actin filaments and an accumulation of filamentous actin at the leading edges in DMSO controls (top). 0136-0228-treated cells (25μM) show diffuse actin staining and collapsed morphology (bottom). **(E)** Microtubule polymerization assay confirming 1327-0036 as a tubulin inhibitor. In this assay purified porcine tubulin is incubated at a high concentration and allowed to polymerize over time in the presence of either the negative control DMSO or varying concentrations of the 1327-0036 inhibitor, using nocodazole as a positive control. Polymerization reactions are monitored as an increase in light scattering over time. DMSO has no inhibitory effect on the polymerization reaction whereas the 1327-0036 inhibitor shows a dose-dependent inhibition of polymerization and potency of near equivalence to 5μM nocodazole. **(F)** Quantitation of the effects of 1327-0036 on yeast cells. Cells were incubated in the presence of compound as described above. One hundred cells for each treatment were visually inspected and scored for three interphase phenotypes (depolymerized, partially depolymerized, normal) and two mitotic phenotypes (normal and abnormal). **(G)** HeLa cells treated with DMSO, nocodazole or 1327-0036 (*10*). DMSO-treated cells (top) show a normal microtubule organizing center (visualized in green via GFP-tubulin) that is abolished in 12.9nM nocodazole (middle) and 30nM 1327-0036 (bottom). Nuclei and chromosomes are stained blue. Mitotic cells show normal metaphase spindles which are, in contrast, fragmented in both nocodazole and 1327-0036 (scale bar = 25μm). **(H)** A fluorescent colchicine binding displacement assay shows that 1327-0036 binds to the same site on the tubulin heterodimer as colchicine. This assay is performed using purified colchicine-tubulin complexes (which are inherently fluorescent) and then measuring the loss of fluorescence in either DMSO (100μM; 100±0.7%) or vinblastine (a vinca alkaloid microtubule depolymerizer that binds to a distinct site on the tubulin heterodimer; 100μM; 93.6±1.3%). The fluorescence is unaffected with either DMSO or vinblastine confirming that vinblastine does not bind to the same site as colchicine. In contrast, competition with the 1327-0036 (100μM; 45.9±1.2%) compound decreases the colchicine tubulin fluorescence by half, presumably by displacing the tubulin-colchicine binding. Data are mean +/- SD (N = 3).

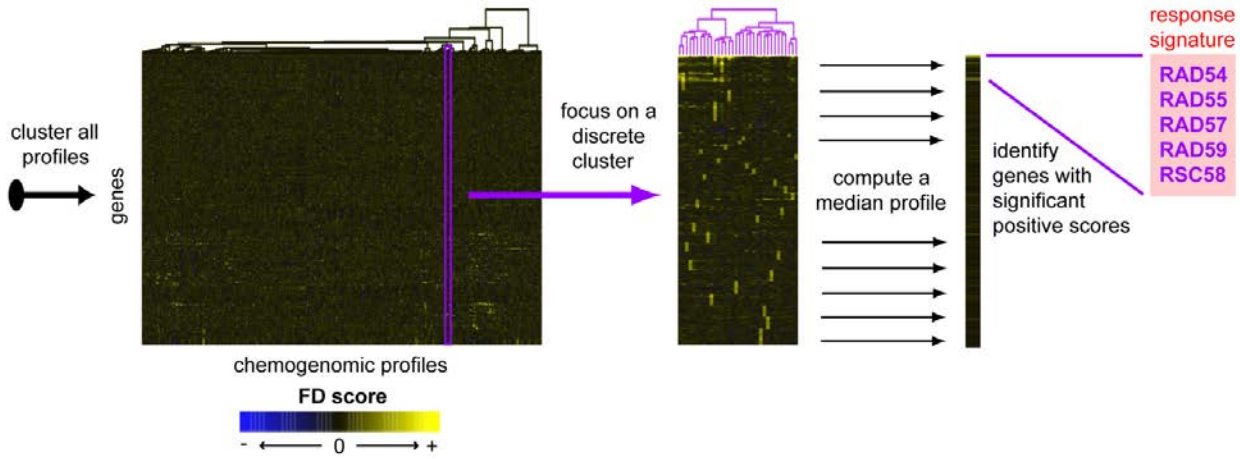


Fig. S11

Definition of response signatures. The steps involved in defining the DNA damage response signature, as an example. First, all HIPHOP profiles are hierarchically clustered. Dynamic branch cutting is then used to identify discrete clusters (*10*). For a given cluster, a median profile is computed. Genes with significant FD scores (standard normal distribution $P < 0.001$) in the median profile define a signature. Therefore, each signature is a set of genes, and is associated with a set of profiles, and by extension, the set of profiled compounds.

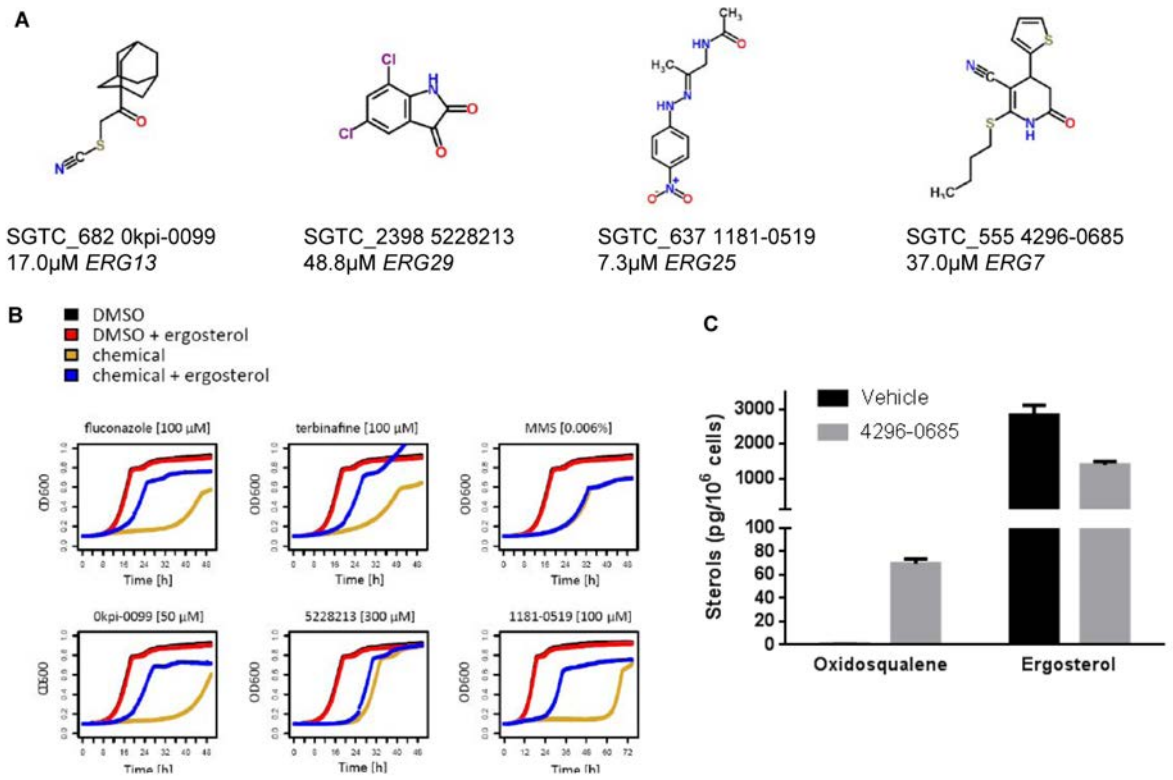


Fig. S12

Validation of Erg13, Erg29, Erg25 and Erg7 inhibitors by ergosterol rescue and LC/MS/MS (Erg7 only). (A) Chemical structures of four compounds predicted by HIP to inhibit proteins in the ergosterol biosynthesis pathway (indicate below the structure). We also show the Screen ID, name, screening concentration of the compounds. **(B)** Validation of Erg13, Erg29 and Erg25 inhibitors by ergosterol rescue. The strain WPY361 (upc2-1), which allows efficient uptake of ergosterol from the culture medium under aerobic conditions (*64*), was grown under four conditions: DMSO (black), DMSO + ergosterol (red), chemical (gold), and chemical + ergosterol (blue). Each plot represents a unique chemical inhibitor (indicated above each plot, with the concentration used in squared brackets). Ergosterol stock (5mM in Tween:ethanol [1:1]) was added for a final concentration of 50µM. Equal volume of Tween:ethanol was added to samples without ergosterol as a control. Methyl methanesulfonate (MMS), a DNA damaging agent, served as a negative control. Top panel: ergosterol rescues the growth inhibition induced by fluconazole and terbinafine, established inhibitors of the ergosterol biosynthesis pathway, but has no effect on the growth inhibition by MMS, a DNA alkylating agent. Bottom panel: ergosterol rescues the growth inhibition of the putative ergosterol biosynthesis inhibitors 0kpi-0099, 5228213 and 1181-0519. Note: *ERG29* did not pass the HIP hit criteria but nonetheless validated. **(C)** Confirmation of a Erg7 inhibitor by liquid chromatography tandem mass spectrometry (LC/MS/MS). Quantification of 2,3-oxidosqualene, the precursor of the lanosterol synthase Erg7, and ergosterol (y-axis) in DMSO- (black) and in 4296-0685-treated (grey) cells as measured by LC/MS. Experiments were performed in triplicate; shown is a representative graph of one of three biological replicate experiments, with 2-3 technical replicates for each experiment; data are mean +/- SD across the technical replicates. These results show that 4296-0685 induced increased oxidosqualine and decreased ergosterol content in cells, consistent with inhibition of Erg7 (*10*).

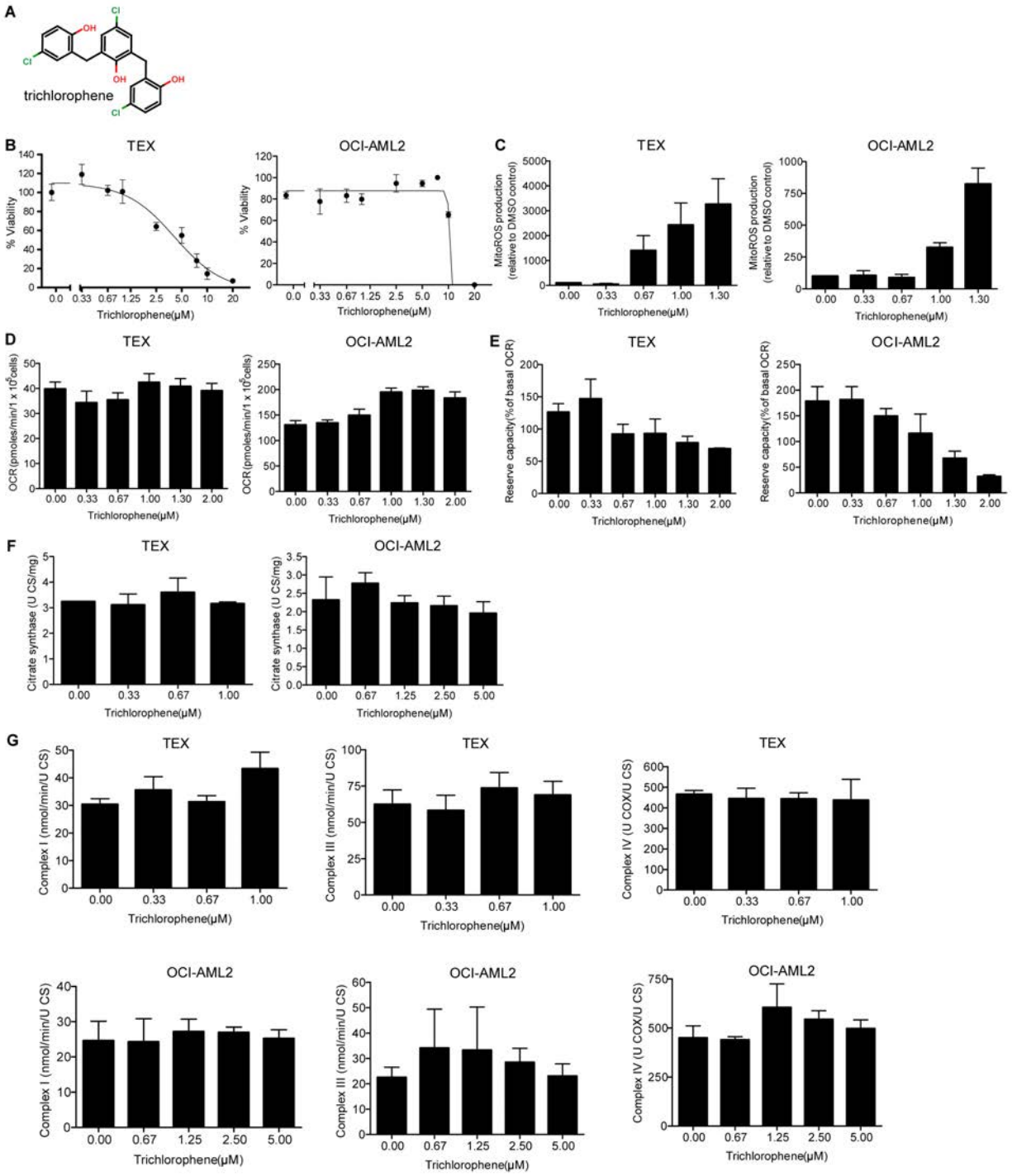


Fig. S13

Mitochondrial activity of trichlorophene in mammalian cells (OCI-AML2 and TEX leukemia cells). (A) Chemical structure of trichlorophene. **(B-G)** Cells were treated with increasing concentrations of trichlorophene for 72h. The data are mean +/- SD (*N* = 3). **(B)** Cell growth and viability were measured by CellTiter-Fluor assay. **(C)** Mitochondrial reactive oxygen species (MitoROS) production was measured using MitoSOX Red staining and flow cytometry. **(D)** 120,000 trypan-blue negative viable cells/well were seeded into 96-well plates and basal oxygen consumption rate (OCR) was measured with the Seahorse Metabolic Flux Analyzer as

described in (*10*). **(E)** Reserve capacity was measured in cells treated as in panel D along with the addition of 1µM of Carbonyl cyanide p-[trifluoromethoxy]-phenyl-hydrazone (FCCP) in TEX cells and 1.5µM FCCP in OCI-AML2 cells. Reserve capacity represents the difference between the minimal and maximal respiratory rates. **(F)** Mitochondrial mass was evaluated by measuring citrate synthase activity. **(G)** Respiratory chain complex I, III, and IV enzyme activity relative to citrate synthase activity was measured as described in (*10*).

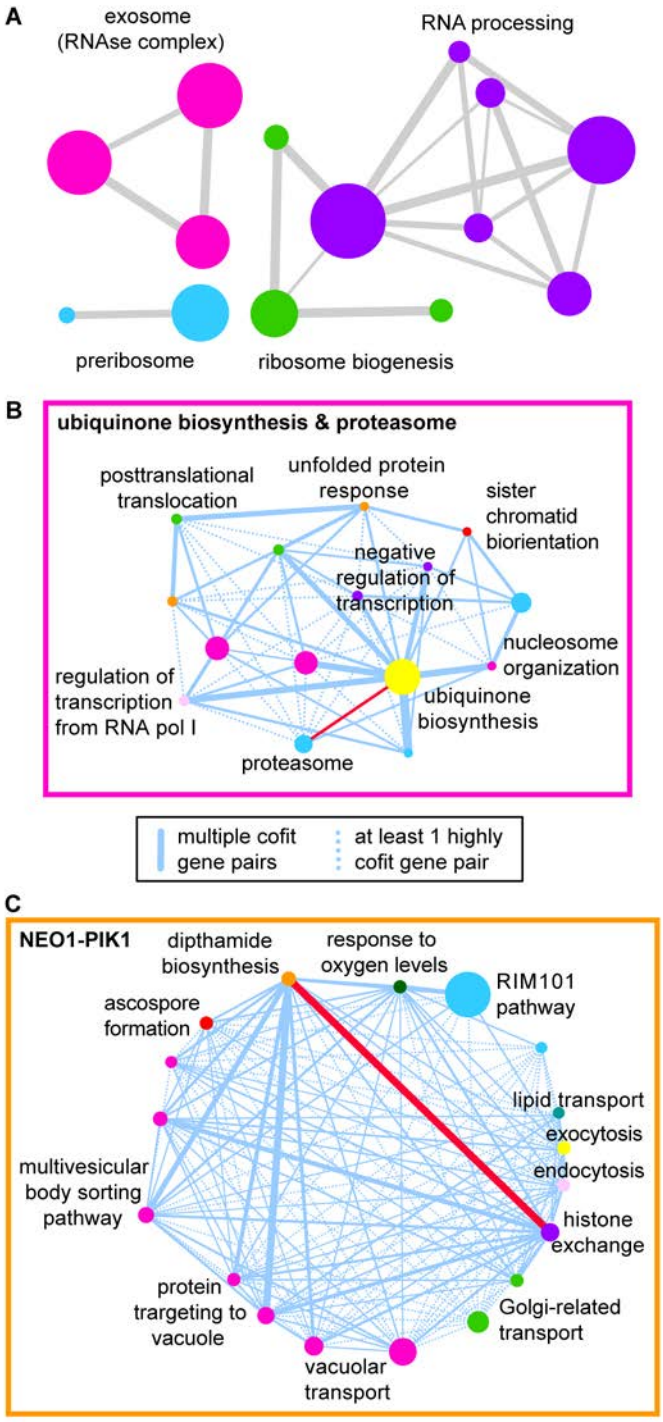


Fig. S14

Response signatures identify relationships between biological processes. (A) Protein complexes and biological processes enriched in the *exosome* signature (visualization as in **fig. S2**). Relationships between biological processes identified by the **(B)** *ubiquinone biosynthesis & proteasome* and **(C)** *NEO1-PIK1* signatures. Each box contains a cofitness network specific to a single signature, where nodes represent processes enriched in the signature (hypergeometric test FDR ≤ 0.1), and edges indicate cofitness [i.e. gene correlation (*10*)] across the HIPHOP profiles associated with the signature. Solid edges indicate that ≥5% of the gene pairs bridging the

processes exhibit significant cofitness (Student's t-distribution method $P \leq 0.01$); width is proportional to the percentage. Dotted edges indicate that ≥1 bridging gene pairs exhibits very strong cofitness (≥ 0.8). Red edges highlight relationships discussed in the text. Node size is proportional to the significance of enrichment; shared color indicates processes with many shared genes.

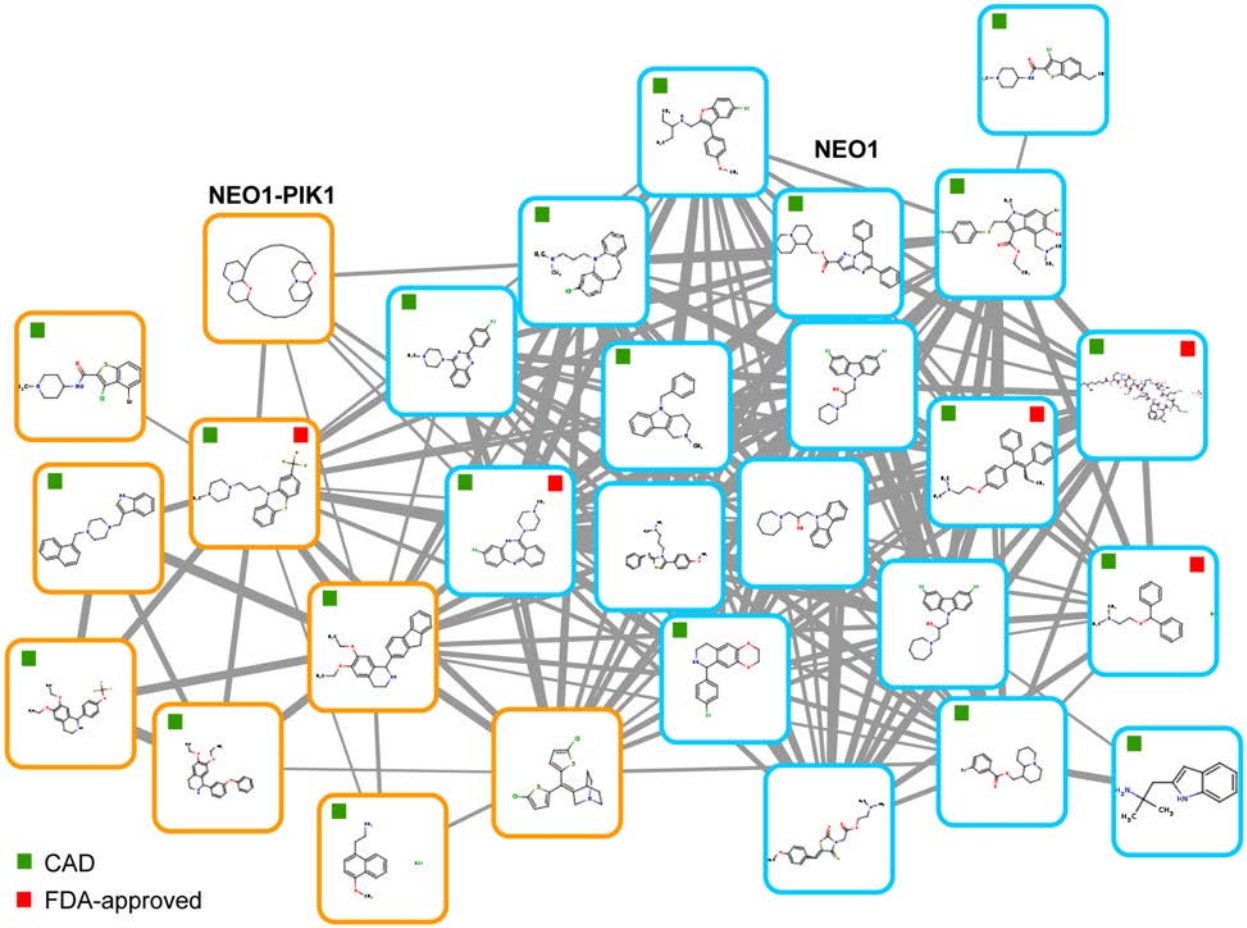


Fig. S15

Coinhibition network of compounds inducing a *NEO1* or *NEO1-PIK1* signature. Nodes represent NEO1 or NEO1-PIK1 compounds that exhibit strong coinhibition (i.e. profile correlation (*10*); coinhibition ≥ 0.5, Student's t-distribution method $P \leq 0.01$); border color indicates signature. Edge width is proportional to the degree of coinhibition. Compounds that are cationic amphiphilic drugs (CADs) and/or FDA-approved are indicated with green and red squares, respectively.

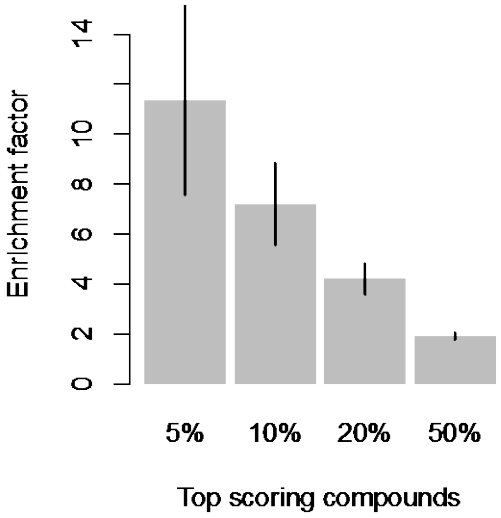


Fig. S16

Cross validation of the *NEO1*-based structural model. We generated a structural model predictive of compounds that induce the *NEO1* or *NEO1-PIK1* signature (i.e. actives), and therefore, top-scoring compounds (determined by the model) are likely to induce a *NEO1*-based response. The bar plot shows the mean +/- SD of the enrichment factor (<number of actives in compound set> / <expected number of actives by random sampling>) for top scoring compounds (e.g. top 5%), obtained by five-fold cross-validation of the model. If we predict that the top 10% are actives, the model performed ~7-fold better than random, and the mean percentage of correct predictions across the cross-validation runs is ~99%.

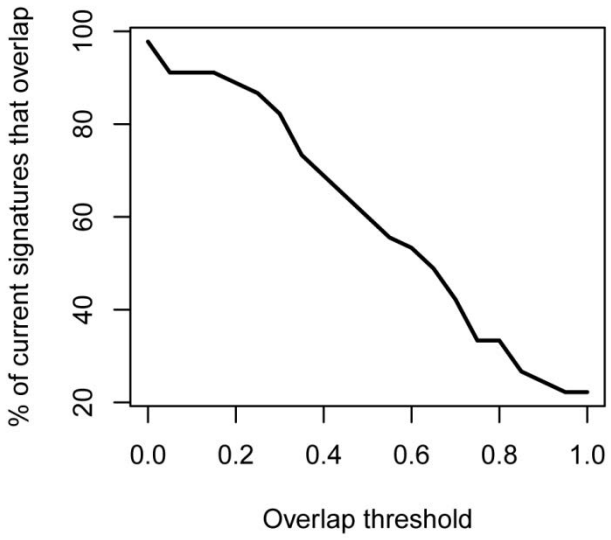


Fig. S17

Major response signatures that are similar to signatures derived from an earlier chemogenomic dataset. The percentage of major response signatures that overlap any signature derived from the Hillenmeyer *et al.* 2008 (*7*) dataset (y-axis), at different overlap thresholds (<size of overlap between signature gene sets> / <size of the smaller of the two signature gene sets>; x-axis). If we require overlap ≥ 0.5, 60% of the signatures derived from the current dataset are similar to a Hillenmeyer signature (**table S4**).

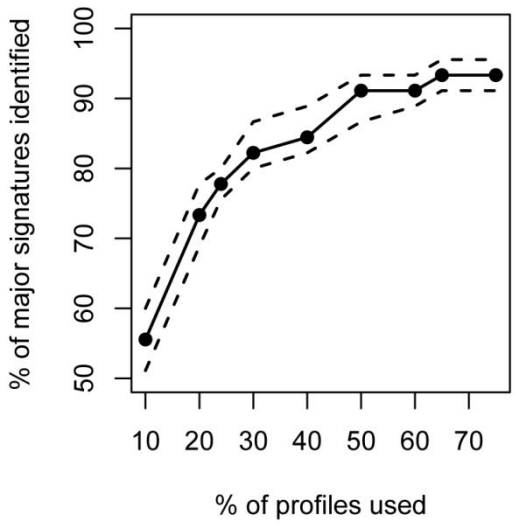


Fig. S18

Percentages of response signatures identified using subsets of HIPHOP profiles. We repeatedly sampled random subsets (of increasing size) of our HIPHOP profiles, clustered them, defined new signatures from these clusters, and determined the percentage of the 45 major signatures recovered by the new signatures (*10*). For example, we sampled 30% of the profiles 100 times, and for each sample, we defined new signatures. The median percentage of major signatures recovered by each set of new signatures is ~82% (see solid line). For each percentage of profiles used, 50% of the recovery percentages fall within the interval specified by the first and third quartiles (dashed lines below and above the solid line, respectively).

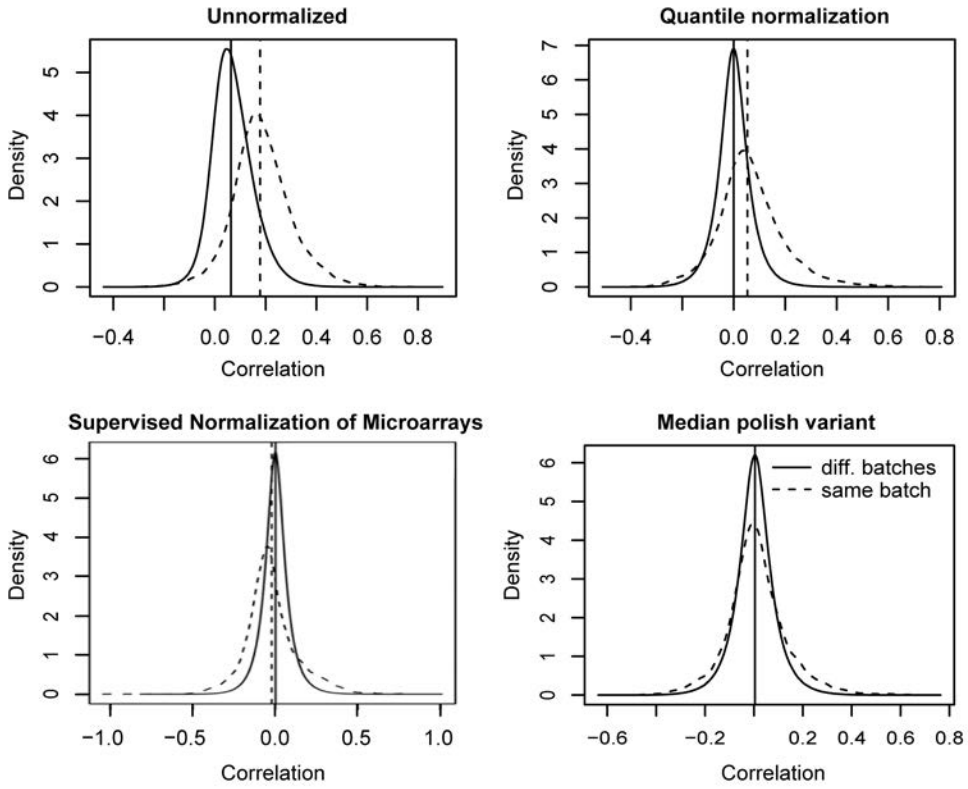


Fig. S19

Detection of batch effects. Each plot show the distributions of profile correlations corresponding to profiles generated in the same chip date batch (i.e. hybridized on the same day; dashed line) or in different batches (solid line), given FD scores computed from unnormalized intensities (top-left), quantile normalized intensities (top-right) and intensities normalized with Supervised Normalization of Microarrays (bottom-left) or the median polish variant (bottom-right) (*10*). For the unnormalized intensities, FD scores were computed from the unnormalized 'best tag' intensities, excluding tags with intensities at background level, and using experimental group-matched controls. In the absence of batch effects, we expect the distribution of correlations (Pearson's correlation coefficient) between profiles in the same batch to be similar to the distribution of correlations between profiles from different batches. The vertical lines indicate the medians of the distributions. For the unnormalized and quantile normalized intensities, profiles in the same batch tend to be more highly correlated than experiments from different batches, thus illustrating batch effects. These batch effects are greatly reduced when using intensities normalized with the median polish variant.

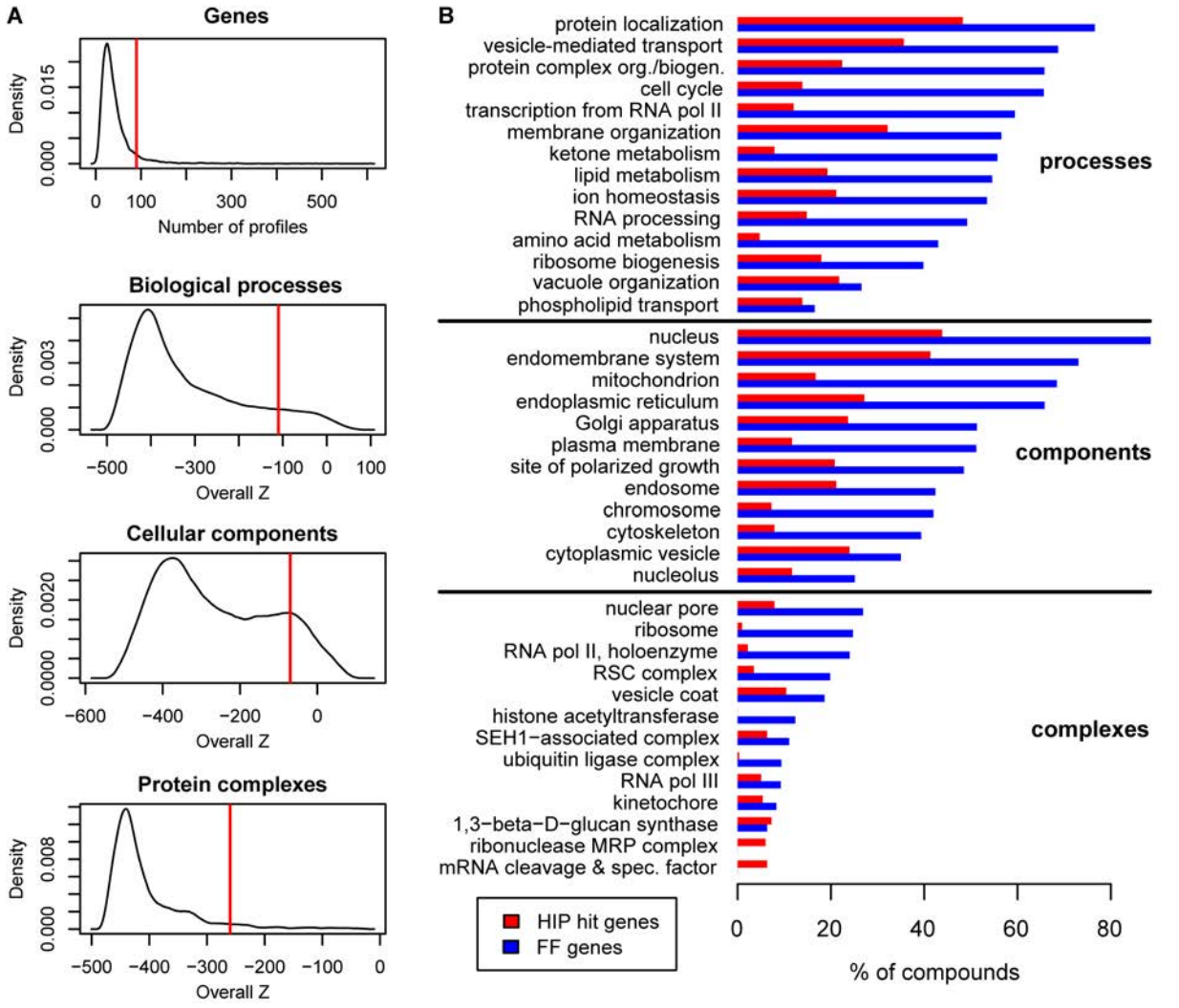


Fig. S20

Genes, biological processes, cellular components and protein complexes frequently affected by compounds. (A) Thresholds used to identify genes, processes, components and complexes frequently affected by compounds. A gene is defined as a frequent flier (FF) if it satisfies two conditions (*10*), one of which requires that the number of profiles in which the gene has a significant positive FD score is greater than or equal to the selected threshold. The top plot shows the number-of-profiles distribution across all genes, and the selected threshold at the bend point (red line). A process/component/complex is defined as a frequent term if its overall *Z* score, measuring the significance with which it is enriched across all profiles, is greater than or equal to the selected threshold. The last three plots show the *Z* distribution of each term class, and the selected thresholds at the bend points (red lines). **(B)** Blue bars indicate the percentage of screened compounds (total = 3,250) that exhibit a chemical-genetic interaction with a FF gene annotated to a given process, component or complex. For example, 77% of screened compounds interact with a FF gene involved in protein localization. Red bars indicate the percentage of compounds with HIP hits (total = 317 compounds) where a HIP hit is annotated to the given gene category. For example, 10% of the compounds have a HIP hit in the vesicle coat (*10*).

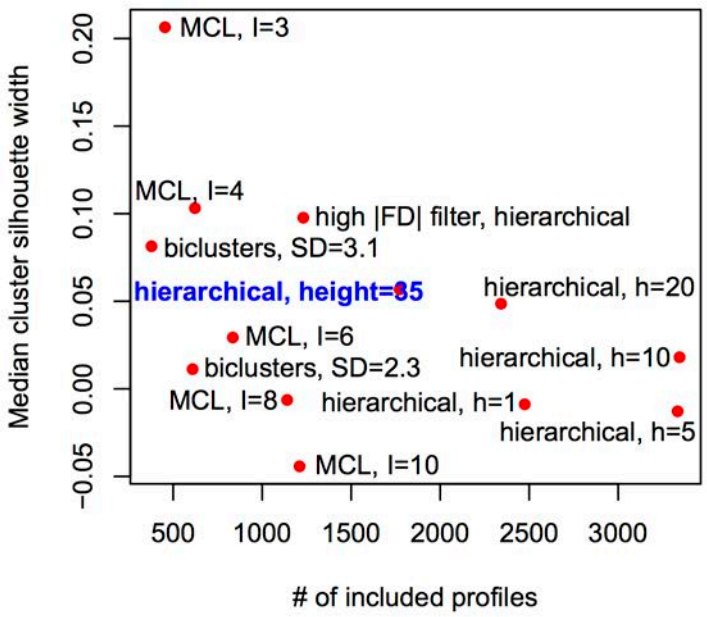


Fig. S21

Evaluation of profile clusters by profile coverage and silhouette statistics. Each point represents a set of HIPHOP profile clusters resulting from a distinct methodology, and distinct parameter settings are indicated. The x-axis indicates the number of profiles included in each set of clusters, i.e. profile coverage. The y-axis indicates the median cluster silhouette width for each set of clusters (*10*), where higher values indicate that profiles were clustered better (i.e. on average, a profile is more similar to profiles in the same cluster, than to profiles in other clusters). Abbreviations: FD – fitness defect score; h – maximum joining height; I – inflation; MCL – Markov clustering; SD – standard deviation threshold.

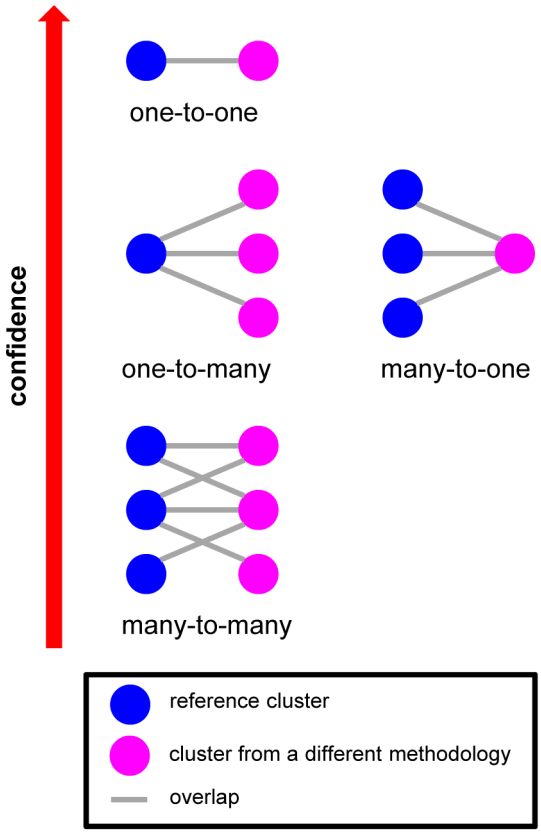


Fig. S22

Mappings between sets of clusters resulting from distinct methodologies that confer confidence in the robustness of clusters. Clusters that are identified by multiple clustering methodologies are robust to the methodology. To identify robust clusters, we designated one set of clusters as the reference, and quantified the overlap between reference clusters (blue nodes) and clusters resulting from a different methodology (pink nodes). Substantial overlaps (i.e. overlap coefficient = <size of cluster overlap> / <size of the smaller of the two clusters> ≥ 0.5) enabled a mapping between sets of clusters form distinct methodologies, and a one-to-one mapping shows that a reference cluster was identified by both methodologies, thereby providing evidence for the robustness of that cluster. The other types of mappings shown also indicate that a reference cluster was identified by both methodologies, although the cluster is represented in a different way, and thus the evidence for robustness is weaker. We used these different types of mappings to assign a level of confidence in the robustness of a given cluster (*10*).

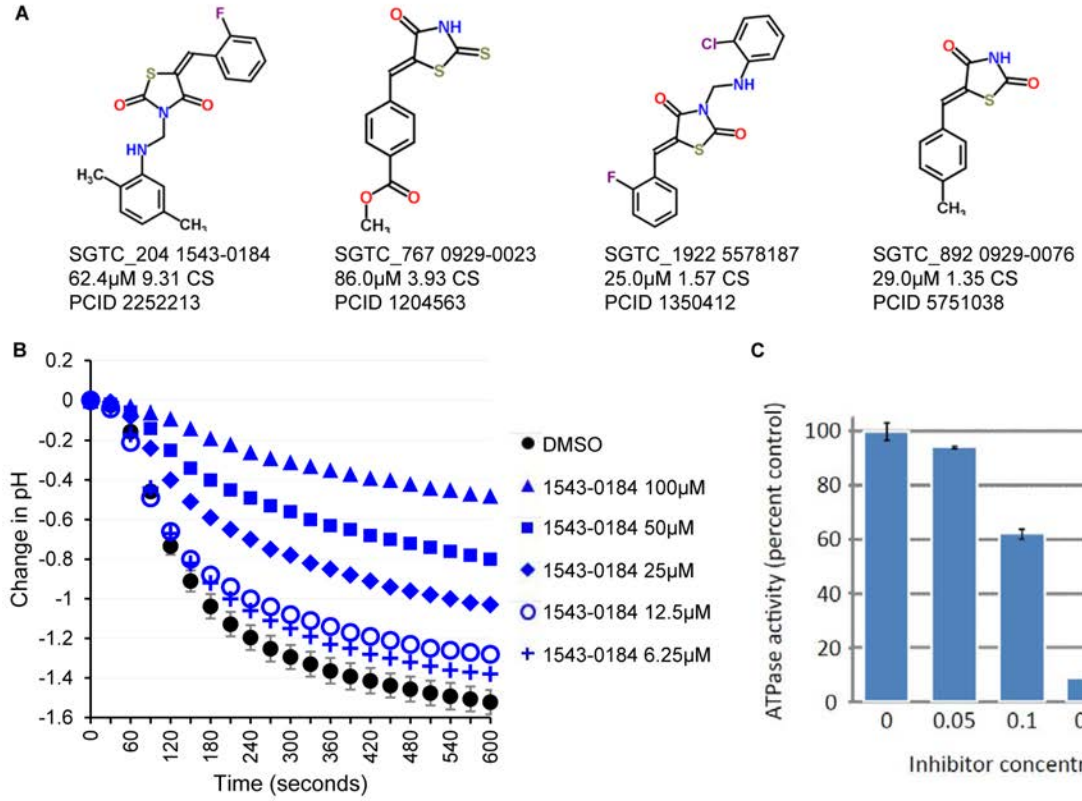


Fig. S23

Validation of the Pma1 inhibitor by *in vitro* assays. (A) Chemical structures of four thiazolidinediones, one of which, 1543-0184, designated *PMA1* as a HIP hit. We also show the Screen ID, name, screening concentration, *PMA1* clearance score (CS) and PubChem CID of the compounds. **(B)** Media acidification by 1543-0184. Glucose-induced media acidification (*65*) was monitored by measuring pH changes (y-axis) over time (x-axis) in the presence of 100µM, 50µM, 25µM, 12.5µM, 6.25µM 1543-0184 or DMSO (control) as indicated. DMSO data are mean +/- SD (N = 3). **(C)** ATPase activity in the presence of 1543-0184. Vanadate-sensitive, ATPase-specific activity was measured from purified yeast plasma membranes in the presence of increasing concentrations of 1543-0184. Percentage of the no-compound control's ATPase activity is indicated for each sample, and data are mean +/- SD (N = 3).

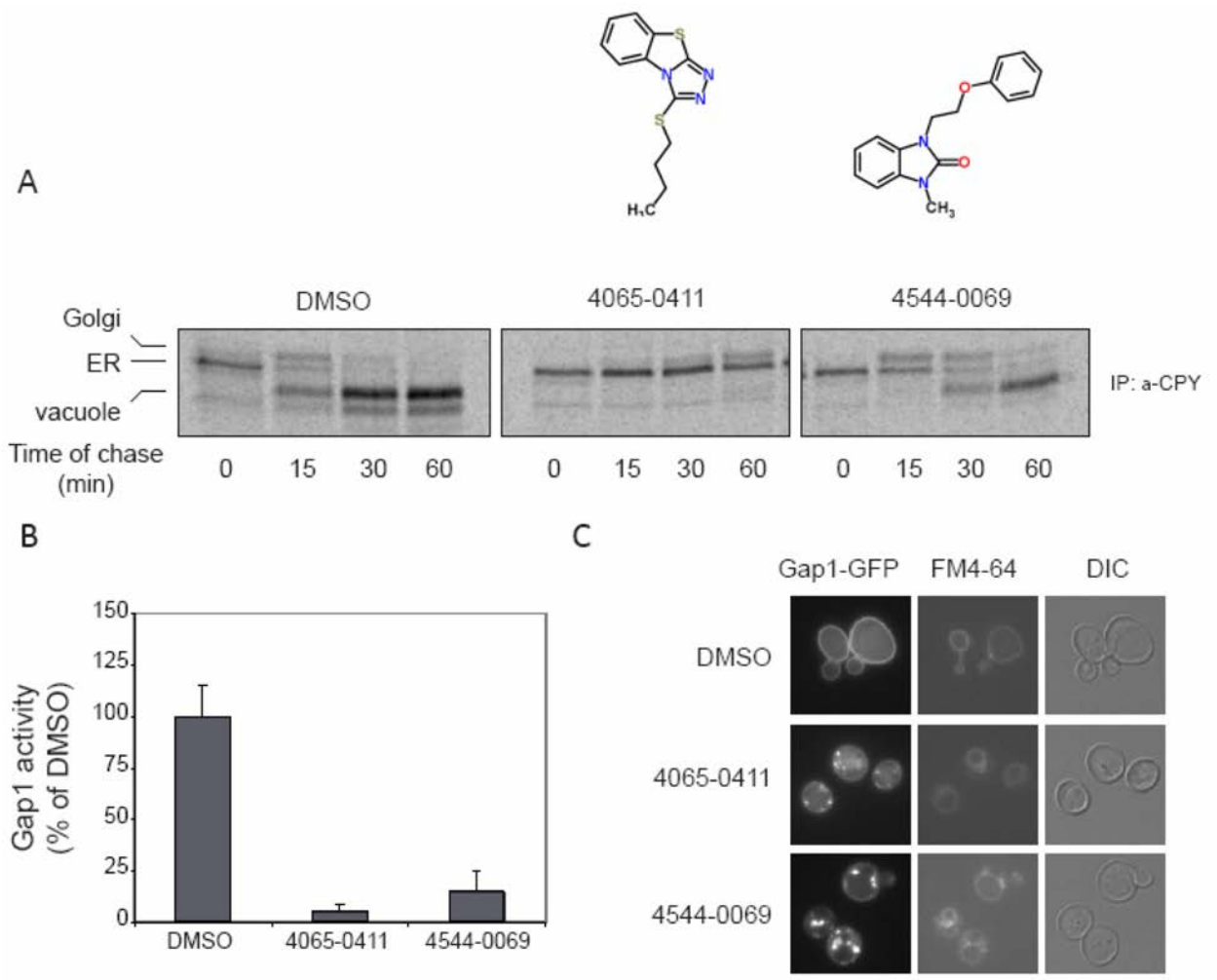


Fig. S24

***In vivo* validation of Sec13 inhibitors. (A)** 4065-0411 and 4544-0069 inhibit carboxypeptidase Y (CPY) processing. Wild-type cells (ESY360) were grown in synthetic medium lacking methionine, treated with DMSO or 0.5mM of compound (labels above image), and then pulse-labeled with 35S-methionine/cysteine. A chase was initiated by addition of cold methionine and cysteine, and aliquots were removed at 15, 30 and 60 minutes (labels below image). 35S-labelled CPY was immunoprecipitated (IP) and then size-fractionated by SDS-PAGE. Endoplasmic reticulum (ER), Golgi and vacuole isoforms of CPY are indicated on the left. **(B)** Gap1 activity is drastically reduced in cells treated with 4065-0411 or 4544-0069, indicative of a block in transport to the plasma membrane due to inhibition of Sec13. Gap1 activity was assessed by measuring the rate of uptake of [14C]-citrulline (*61*). Data are mean +/- SD ($N = 3$). **(C)** Gap1-GFP localization is perturbed in cells treated with 4065-0411 or 4544-0069. Gap1-GFP expressing cells were grown in SD ammonium and labeled with FM4-64 (which labels endosomes and vacuoles). Cells were treated with DMSO, 4065-0411, or 4544-0069 (indicated on the left), fixed, and imaged with a fluorescent microscope. DIC: differential interference contrast. Note: Though *SEC13* was not a designated HIP hit in the 4065-0411 screen, it nonetheless was confirmed as a likely target in the assays described above.

Additional Data table S1 (separate file)

Compound library. A list of screens, with information on the compounds screened. See ReadMe sheet for details.

Additional Data table S2 (separate file)

HIP hits. A list of HIP hits (*10*), additional compound-gene interactions supported by follow-up experiments, and established compound-target interactions. See ReadMe sheet for details.

Additional Data table S3 (separate file)

Inhibitory concentrations of 152 compounds for yeast versus mammalian cells. The curves fitted to the dose-response data in **fig. S4** were used to estimate the compound concentrations required to inhibit growth of yeast and mammalian cells by 20% and 50% (i.e. IC_{20} and IC_{50}). The yeast ICs are based on the growth of the wildtype BY4743 strain. The mammalian ICs are based on the growth of one or more of five cell lines (A549, AML2, HCT116, LP1, TEX). All ICs are expressed in μM. See ReadMe sheet for details.

Additional Data table S4 (separate file)

Response signatures. The genes defining major and minor response signatures. An indication of whether or not each major signature was detected in our prior chemogenomic dataset (*10*) is also provided. See ReadMe sheet for details.

Additional Data table S5 (separate file)

Enrichment of signature compounds with compounds known to have similar bioactivity. Various chemical ontologies were used to identify classes of compounds known to have similar bioactivity, where some classes indicate shared general bioactivity (e.g. antifungal activity) and others indicate shared mechanism of action. In addition to links between major response signatures and bioactivities identified by enrichment (hypergeometric text FDR < 0.1), the bioactivities of uncharacterized compounds inferred by guilt-by-signature-association are provided. Specifically, for each of the 11 signatures associated with ≥2 compounds known to have similar mechanism, we list the uncharacterized compounds associated with those signatures and infer that their bioactivities are similar to those of the characterized compounds associated with the same signature. Five of those signatures are linked to bioactivities by enrichment, and thus, the bioactivity inferences involving these signatures have more supporting evidence. See ReadMe sheet for details.

Additional Data table S6 (separate file)

Enrichment of genetic interaction profiles with response signatures. Of the genetic interaction profiles that are not significantly enriched for any Gene Ontology (GO) category, the genetic profiles that are significantly enriched for major signatures are provided. Specifically, we used each genetic profile to identify the set of genes that negatively genetically interact with the associated query gene (*10*). The resulting gene set was then tested for enrichment with major response signatures, and significant enrichment (hypergeometric test FDR < 0.1) is reported here. See ReadMe sheet for details.

Additional Data table S7 (separate file)

Process relationships. Relationships between biological processes, shown in red in **fig. S14B,C**, are supported by cofitness interactions between genes, where one gene belongs to one process and the second gene to the second process. Moreover, these cofitness interactions were identified in specific major response signatures. Where applicable, the table also includes genetic and physical interactions (obtained from BioGRID) supporting the listed relationships. See ReadMe sheet for details.

Additional Data table S8 (separate file)

Enrichment of signature compounds with chemical moieties. Major signatures representing compounds enriched for specific chemical moieties are provided. Briefly, structural moieties/fragments were identified in the compounds screened. For each major signature, we assessed whether the associated compounds are enriched for specific fragments. Here we show the signatures that are significantly enriched for some fragment (hypergeometric test FDR < 0.1), where the fraction of signature compounds containing the fragment is greater than > 0.2. See ReadMe sheet for details.

Additional Data table S9 (separate file)

Strain list. A list of deletion strains used in this study. Fitness defects were measured for each strain. See ReadMe sheet for details.

Additional Data table S10 (separate file)

Robustness of the HIPHOP profile clusters. The HIPHOP profile clusters used to define the major response signatures and the levels of confidence in their robustness. Briefly, hierarchical clusters (no profile filter) were used as a reference set of clusters, and each of these clusters was tested for overlap with clusters generated by three other methodologies. The identification of reference clusters by other methodologies gives confidence in the robustness of those clusters (*10*). See ReadMe sheet for details.

References and Notes

1. P. M. Barrett, A. Alagely, E. J. Topol, Cystic fibrosis in an era of genomically guided therapy. *Hum. Mol. Genet.* **21** (R1), R66–R71 (2012). Medline doi:10.1093/hmg/dds345

2. J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J. P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, T. R. Golub, The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006). Medline doi:10.1126/science.1132939

3. F. Iorio, R. Bosotti, E. Scacheri, V. Belcastro, P. Mithbaokar, R. Ferriero, L. Murino, R. Tagliaferri, N. Brunetti-Pierri, A. Isacchi, D. di Bernardo, Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14621–14626 (2010). Medline doi:10.1073/pnas.1000138107

4. G. Giaever, A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Véronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. André, A. P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Güldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kötter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis, M. Johnston, Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002). Medline doi:10.1038/nature00935

5. G. Giaever, P. Flaherty, J. Kumm, M. Proctor, C. Nislow, D. F. Jaramillo, A. M. Chu, M. I. Jordan, A. P. Arkin, R. W. Davis, Chemogenomic profiling: Identifying the functional interactions of small molecules in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 793–798 (2004). Medline doi:10.1073/pnas.0307490100

6. P. Y. Lum, C. D. Armour, S. B. Stepaniants, G. Cavet, M. K. Wolf, J. S. Butler, J. C. Hinshaw, P. Garnier, G. D. Prestwich, A. Leonardson, P. Garrett-Engele, C. M. Rush, M. Bard, G. Schimmack, J. W. Phillips, C. J. Roberts, D. D. Shoemaker, Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* **116**, 121–137 (2004). Medline doi:10.1016/S0092-8674(03)01035-3

7. M. E. Hillenmeyer, E. Fung, J. Wildenhain, S. E. Pierce, S. Hoon, W. Lee, M. Proctor, R. P. St.Onge, M. Tyers, D. Koller, R. B. Altman, R. W. Davis, C. Nislow, G. Giaever, The chemical genomic portrait of yeast: Uncovering a phenotype for all genes. *Science* **320**, 362–365 (2008). Medline doi:10.1126/science.1150021

8. G. Giaever, D. D. Shoemaker, T. W. Jones, H. Liang, E. A. Winzeler, A. Astromoff, R. W. Davis, Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat. Genet.* **21**, 278–283 (1999). Medline doi:10.1038/6791

9. A. B. Parsons, R. L. Brost, H. Ding, Z. Li, C. Zhang, B. Sheikh, G. W. Brown, P. M. Kane, T. R. Hughes, C. Boone, Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat. Biotechnol.* **22**, 62–69 (2004). Medline doi:10.1038/nbt919

10. Materials and methods are available as supplementary material on *Science* Online.

11. I. M. Wallace, M. L. Urbanus, G. M. Luciani, A. R. Burns, M. K. Han, H. Wang, K. Arora, L. E. Heisler, M. Proctor, R. P. St.Onge, T. Roemer, P. J. Roy, C. L. Cummins, G. D. Bader, C. Nislow, G. Giaever, Compound prioritization methods increase rates of chemical probe discovery in model organisms. *Chem. Biol.* **18**, 1273–1283 (2011). Medline doi:10.1016/j.chembiol.2011.07.018

12. N. P. Torres, A. Y. Lee, G. Giaever, C. Nislow, G. W. Brown, A high-throughput yeast assay identifies synergistic drug combinations. *Assay Drug Dev. Technol.* **11**, 299–307 (2013). Medline doi:10.1089/adt.2012.503

13. Q. Hu, W. J. Nelson, E. T. Spiliotis, Forchlorfenuron alters mammalian septin assembly, organization, and dynamics. *J. Biol. Chem.* **283**, 29563–29571 (2008). Medline doi:10.1074/jbc.M804962200

<jrn>14. A. H. Nile, A. Tripathi, P. Yuan, C. J. Mousley, S. Suresh, I. M. Wallace, S. D. Shah, D. T. Pohlhaus, B. Temple, C. Nislow, G. Giaever, A. Tropsha, R. W. Davis, R. P. St.Onge, V. A. Bankaitis, PITPs as targets for selectively interfering with phosphoinositide signaling in cells. *Nat. Chem. Biol.* **10**, 76–84 (2014). Medline doi:10.1038/nchembio.1389</jrn>

15. A. E. Cleves, T. P. McGee, E. A. Whitters, K. M. Champion, J. R. Aitken, W. Dowhan, M. Goebl, V. A. Bankaitis, Mutations in the CDP-choline pathway for phospholipid biosynthesis bypass the requirement for an essential phospholipid transfer protein. *Cell* **64**, 789–800 (1991). Medline doi:10.1016/0092-8674(91)90508-V

16. M. Fang, B. G. Kearns, A. Gedvilaite, S. Kagiwada, M. Kearns, M. K. Fung, V. A. Bankaitis, Kes1p shares homology with human oxysterol binding protein and participates in a novel regulatory pathway for yeast Golgi-derived transport vesicle biogenesis. *EMBO J.* **15**, 6447–6459 (1996). Medline

17. A. Chatr-aryamontri, B. J. Breitkreutz, S. Heinicke, L. Boucher, A. Winter, C. Stark, J. Nixon, L. Ramage, N. Kolas, L. O'Donnell, T. Reguly, A. Breitkreutz, A. Sellam, D. Chen, C. Chang, J. Rust, M. Livstone, R. Oughtred, K. Dolinski, M. Tyers, The BioGRID interaction database: 2013 update. *Nucleic Acids Res.* **41**, D816–D823 (2013). Medline doi:10.1093/nar/gks1158

18. M. Costanzo, A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Koh, K. Toufighi, S. Mostafavi, J. Prinz, R. P. St.Onge, B. VanderSluis, T. Makhnevych, F. J. Vizeacoumar, S. Alizadeh, S. Bahr, R. L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z. Y. Lin, W. Liang, M. Marback, J. Paw, B. J. San Luis, E. Shuteriqi, A. H. Tong, N. van Dyk, I. M. Wallace, J. A. Whitney, M. T. Weirauch, G. Zhong, H. Zhu, W. A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pál, F. P. Roth, G. Giaever, C. Nislow, O. G. Troyanskaya, H. Bussey, G. D. Bader, A. C. Gingras, Q. D. Morris, P. M. Kim, C. A. Kaiser, C. L. Myers, B. J. Andrews, C. Boone, The genetic landscape of a cell. *Science* **327**, 425–431 (2010). Medline doi:10.1126/science.1180823

19. N. Anderson, J. Borlak, Drug-induced phospholipidosis. *FEBS Lett.* **580**, 5533–5540 (2006). Medline doi:10.1016/j.febslet.2006.08.061

20. M. J. Reasor, S. Kacew, Drug-induced phospholipidosis: Are there functional consequences? *Exp. Biol. Med. (Maywood)* **226**, 825–830 (2001). Medline

21. M. M. Rainey, D. Korostyshevsky, S. Lee, E. O. Perlstein, The antidepressant sertraline targets intracellular vesiculogenic membranes in yeast. *Genetics* **185**, 1221–1233 (2010). Medline doi:10.1534/genetics.110.117846

22. M. Škrtić, S. Sriskanthadevan, B. Jhas, M. Gebbia, X. Wang, Z. Wang, R. Hurren, Y. Jitkova, M. Gronda, N. Maclean, C. K. Lai, Y. Eberhard, J. Bartoszko, P. Spagnuolo, A. C. Rutledge, A. Datti, T. Ketela, J. Moffat, B. H. Robinson, J. H. Cameron, J. Wrana, C. J. Eaves, M. D. Minden, J. C. Wang, J. E. Dick, K. Humphries, C. Nislow, G. Giaever, A. D. Schimmer, Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* **20**, 674–688 (2011). Medline doi:10.1016/j.ccr.2011.10.015

23. R. K. Blackman, K. Cheung-Ong, M. Gebbia, D. A. Proia, S. He, J. Kepros, A. Jonneaux, P. Marchetti, J. Kluza, P. E. Rao, Y. Wada, G. Giaever, C. Nislow, Mitochondrial electron transport is the cellular target of the oncology drug elesclomol. *PLOS ONE* **7**, e29798 (2012). Medline doi:10.1371/journal.pone.0029798

24. in *R Foundation for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria), vol. 2005.

25. S. Hoon, A. M. Smith, I. M. Wallace, S. Suresh, M. Miranda, E. Fung, M. Proctor, K. M. Shokat, C. Zhang, R. W. Davis, G. Giaever, R. P. St.Onge, C. Nislow, An integrated platform of genomic assays reveals small-molecule bioactivities. *Nat. Chem. Biol.* **4**, 498–506 (2008). Medline doi:10.1038/nchembio.100

26. S. E. Pierce, R. W. Davis, C. Nislow, G. Giaever, Genome-wide analysis of barcoded Saccharomyces cerevisiae gene-deletion mutants in pooled cultures. *Nat. Protoc.* **2**, 2958–2974 (2007). Medline doi:10.1038/nprot.2007.427

27. V. Vichai, K. Kirtikara, Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat. Protoc.* **1**, 1112–1116 (2006). Medline doi:10.1038/nprot.2006.179

28. S. Kittanakom *et al.*, Miniature short hairpin RNA screens to characterize antiproliferative drugs. *G3 (Bethesda)* **3**, 1375 (2013).

29. S. E. Pierce, E. L. Fung, D. F. Jaramillo, A. M. Chu, R. W. Davis, C. Nislow, G. Giaever, A unique and universal molecular barcode array. *Nat. Methods* **3**, 601–603 (2006). Medline doi:10.1038/nmeth905

30. S. Jain, G. D. Bader, An improved method for scoring protein-protein interactions using semantic similarity within the gene ontology. *BMC Bioinformatics* **11**, 562 (2010). Medline doi:10.1186/1471-2105-11-562

31. B. H. Mecham, P. S. Nelson, J. D. Storey, Supervised normalization of microarrays. *Bioinformatics* **26**, 1308–1315 (2010). Medline doi:10.1093/bioinformatics/btq118

32. J. J. Benschop, N. Brabers, D. van Leenen, L. V. Bakker, H. W. van Deutekom, N. L. van Berkum, E. Apweiler, P. Lijnzaad, F. C. Holstege, P. Kemmeren, A consensus of core

protein complex compositions for *Saccharomyces cerevisiae*. *Mol. Cell* **38**, 916–928 (2010). Medline doi:10.1016/j.molcel.2010.06.002

33. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).

34. A. C. Zambon, S. Gaj, I. Ho, K. Hanspers, K. Vranizan, C. T. Evelo, B. R. Conklin, A. R. Pico, N. Salomonis, GO-Elite: A flexible solution for pathway and ontology over-representation. *Bioinformatics* **28**, 2209–2210 (2012). Medline doi:10.1093/bioinformatics/bts366

35. D. Merico, R. Isserlin, O. Stueker, A. Emili, G. D. Bader, Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLOS ONE* **5**, e13984 (2010). Medline doi:10.1371/journal.pone.0013984

36. M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, T. Ideker, Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011). Medline doi:10.1093/bioinformatics/btq675

37. M. C. Whitlock, Combining probability from independent tests: The weighted Z-method is superior to Fisher's approach. *J. Evol. Biol.* **18**, 1368–1373 (2005). Medline doi:10.1111/j.1420-9101.2005.00917.x

38. G. W. Bemis, M. A. Murcko, The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **39**, 2887–2893 (1996). Medline doi:10.1021/jm9602928

39. W. H. Halliwell, Cationic amphiphilic drug-induced phospholipidosis. *Toxicol. Pathol.* **25**, 53–60 (1997). Medline doi:10.1177/019262339702500111

40. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, Open Babel: An open chemical toolbox. *J. Cheminform.* **3**, 33 (2011). Medline doi:10.1186/1758-2946-3-33

41. P. Langfelder, B. Zhang, S. Horvath, Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719–720 (2008). Medline doi:10.1093/bioinformatics/btm563

42. S. Van Dongen, Graph clustering via a discrete uncoupling process. *SIAM J. Matrix Anal. Appl.* **30**, 121–141 (2008). doi:10.1137/040608635

43. S. Bergmann, J. Ihmels, N. Barkai, Iterative signature algorithm for the analysis of large-scale gene expression data. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **67**, 031902 (2003). Medline doi:10.1103/PhysRevE.67.031902

44. D. Rogers, R. D. Brown, M. Hahn, Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up. *J. Biomol. Screen.* **10**, 682–686 (2005). Medline doi:10.1177/1087057105281365

45. M. P. Gleeson, A. Hersey, D. Montanari, J. Overington, Probing the links between in vitro potency, ADMET and physicochemical parameters. *Nat. Rev. Drug Discov.* **10**, 197–208 (2011). Medline doi:10.1038/nrd3367

46. A. Goffeau, J. P. Dufour, Plasma membrane ATPase from the yeast *Saccharomyces cerevisiae*. *Methods Enzymol.* **157**, 528–533 (1988). Medline doi:10.1016/0076-6879(88)57101-X

47. G. Schaaf, E. A. Ortlund, K. R. Tyeryar, C. J. Mousley, K. E. Ile, T. A. Garrett, J. Ren, M. J. Woolls, C. R. Raetz, M. R. Redinbo, V. A. Bankaitis, Functional anatomy of phospholipid binding and regulation of phosphoinositide homeostasis by proteins of the sec14 superfamily. *Mol. Cell* **29**, 191–206 (2008). Medline doi:10.1016/j.molcel.2007.11.026

48. G. Schaaf, M. Dynowski, C. J. Mousley, S. D. Shah, P. Yuan, E. M. Winklbauer, M. K. de Campos, K. Trettin, M. C. Quinones, T. I. Smirnova, L. L. Yanagisawa, E. A. Ortlund, V. A. Bankaitis, Resurrection of a functional phosphatidylinositol transfer protein from a pseudo-Sec14 scaffold by directed evolution. *Mol. Biol. Cell* **22**, 892–905 (2011). Medline doi:10.1091/mbc.E10-11-0903

49. S. E. Phillips, B. Sha, L. Topalof, Z. Xie, J. G. Alb, V. A. Klenchin, P. Swigart, S. Cockcroft, T. F. Martin, M. Luo, V. A. Bankaitis, Yeast Sec14p deficient in phosphatidylinositol transfer activity is functional in vivo. *Mol. Cell* **4**, 187–197 (1999). Medline doi:10.1016/S1097-2765(00)80366-4

50. X. Li, S. M. Routt, Z. Xie, X. Cui, M. Fang, M. A. Kearns, M. Bard, D. R. Kirsch, V. A. Bankaitis, Identification of a novel family of nonclassic yeast phosphatidylinositol transfer proteins whose function modulates phospholipase D activity and Sec14p-independent cell growth. *Mol. Biol. Cell* **11**, 1989–2005 (2000). Medline doi:10.1091/mbc.11.6.1989

51. P. A. Spagnuolo, J. Hu, R. Hurren, X. Wang, M. Gronda, M. A. Sukhai, A. Di Meo, J. Boss, I. Ashali, R. Beheshti Zavareh, N. Fine, C. D. Simpson, S. Sharmeen, R. Rottapel, A. D. Schimmer, The antihelmintic flubendazole inhibits microtubule function through a mechanism distinct from Vinca alkaloids and displays preclinical activity in leukemia and myeloma. *Blood* **115**, 4824–4833 (2010). Medline doi:10.1182/blood-2009-09-243055

52. S. Lawo, M. Bashkurov, M. Mullin, M. G. Ferreria, R. Kittler, B. Habermann, A. Tagliaferro, I. Poser, J. R. Hutchins, B. Hegemann, D. Pinchev, F. Buchholz, J. M. Peters, A. A. Hyman, A. C. Gingras, L. Pelletier, HAUS, the 8-subunit human Augmin complex, regulates centrosome and spindle integrity. *Curr. Biol.* **19**, 816–826 (2009). Medline doi:10.1016/j.cub.2009.04.033

53. H. Bénédetti, S. Raths, F. Crausaz, H. Riezman, The END3 gene encodes a protein that is required for the internalization step of endocytosis and for actin cytoskeleton organization in yeast. *Mol. Biol. Cell* **5**, 1023–1037 (1994). Medline doi:10.1091/mbc.5.9.1023

54. A. J. Janssen, F. J. Trijbels, R. C. Sengers, J. A. Smeitink, L. P. van den Heuvel, L. T. Wintjes, B. J. Stoltenborg-Hogenkamp, R. J. Rodenburg, Spectrophotometric assay for complex I of the respiratory chain in tissue samples and cultured fibroblasts. *Clin. Chem.* **53**, 729–734 (2007). Medline doi:10.1373/clinchem.2006.078873

55. M. A. Birch-Machin, I. M. Shepherd, N. J. Watmough, H. S. Sherratt, K. Bartlett, V. M. Darley-Usmar, D. W. Milligan, R. J. Welch, A. Aynsley-Green, D. M. Turnbull, Fatal

lactic acidosis in infancy with a defect of complex III of the respiratory chain. *Pediatr. Res.* **25**, 553–559 (1989). Medline doi:10.1203/00006450-198905000-00025

56. S. Krähenbühl, C. Talos, U. Wiesmann, C. L. Hoppel, Development and evaluation of a spectrophotometric assay for complex III in isolated mitochondria, tissues and fibroblasts from rats and humans. *Clin. Chim. Acta* **230**, 177–187 (1994). Medline doi:10.1016/0009-8981(94)90270-4

57. I. A. Trounce, Y. L. Kim, A. S. Jun, D. C. Wallace, Assessment of mitochondrial oxidative phosphorylation in patient muscle biopsies, lymphoblasts, and transmitochondrial cell lines. *Methods Enzymol.* **264**, 484–509 (1996). Medline doi:10.1016/S0076-6879(96)64044-0

58. X. Zheng, J. M. Shoffner, M. T. Lott, A. S. Voljavec, N. S. Krawiecki, K. Winn, D. C. Wallace, Evidence in a lethal infantile mitochondrial disease for a nuclear mutation affecting respiratory complexes I and IV. *Neurology* **39**, 1203–1209 (1989). Medline doi:10.1212/WNL.39.9.1203

59. A. San Pietro, N. O. Kaplan, S. P. Colowick, Pyridine nucleotide transhydrogenase. VI. Mechanism and stereospecificity of the reaction in *Pseudomonas fluorescens*. *J. Biol. Chem.* **212**, 941–952 (1955). Medline

60. S. Vashist, W. Kim, W. J. Belden, E. D. Spear, C. Barlowe, D. T. Ng, Distinct retrieval and retention mechanisms are required for the quality control of endoplasmic reticulum protein folding. *J. Cell Biol.* **155**, 355–368 (2001). Medline doi:10.1083/jcb.200106123

61. K. J. Roberg, N. Rowley, C. A. Kaiser, Physiological regulation of membrane protein sorting late in the secretory pathway of *Saccharomyces cerevisiae*. *J. Cell Biol.* **137**, 1469–1482 (1997). Medline doi:10.1083/jcb.137.7.1469

62. A. L. Risinger, N. E. Cain, E. J. Chen, C. A. Kaiser, Activity-dependent reversible inactivation of the general amino acid permease. *Mol. Biol. Cell* **17**, 4411–4419 (2006). Medline doi:10.1091/mbc.E06-06-0506

63. W. P. Walters, M. A. Murcko, Prediction of 'drug-likeness'. *Adv. Drug Deliv. Rev.* **54**, 255–271 (2002). Medline doi:10.1016/S0169-409X(02)00003-0

64. Y. Q. Zhang, S. Gamarra, G. Garcia-Effron, S. Park, D. S. Perlin, R. Rao, Requirement for ergosterol in V-ATPase function underlies antifungal activity of azole drugs. *PLOS Pathog.* **6**, e1000939 (2010). Medline doi:10.1371/journal.ppat.1000939

65. D. S. Perlin, C. L. Brown, J. E. Haber, Membrane potential defect in hygromycin B-resistant pma1 mutants of Saccharomyces cerevisiae. *J. Biol. Chem.* **263**, 18118–18122 (1988). Medline