NEOPLASIA
www.neoplasia.com

Volume 13 Number 1 January 2011 pp. 72–80 72

Use of Yeast Chemigenomics and COXEN Informatics in Preclinical Evaluation of Anticancer Agents[1,2]

Steven C. Smith*, Dmytro M. Havaleshko*, Kihyuck Moon*, Alexander S. Baras†, Jae Lee‡, Stefan Bekiranov§, Daniel J. Burke§ and Dan Theodorescu¶

*Department of Urology, University of Virginia, Charlottesville, VA, USA; †Department of Pathology, University of Virginia, Charlottesville, VA, USA; ‡Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA; §Department of Biochemistry & Molecular Genetics, University of Virginia, Charlottesville, VA, USA; ¶University of Colorado Comprehensive Cancer Center, Aurora, CO, USA

Abstract

Bladder cancer metastasis is virtually incurable with current platinum-based chemotherapy. We used the novel COXEN informatic approach for *in silico* drug discovery and identified NSC-637993 and NSC-645809 (C1311), both imidazoacridinones, as agents with high-predicted activity in human bladder cancer. Because even highly effective monotherapy is unlikely to cure most patients with metastasis and NSC-645809 is undergoing clinical trials in other tumor types, we sought to develop the basis for use of C1311 in rational combination with other agents in bladder cancer. Here, we demonstrate in 40 human bladder cancer cells that the *in vitro* cytotoxicity profile for C1311 correlates with that of NSC-637993 and compares favorably to that of standard of care chemotherapeutics. Using genome-wide patterns of synthetic lethality of C1311 with open reading frame knockouts in budding yeast, we determined that combining C1311 with a taxane could provide mechanistically rational combinations. To determine the preclinical relevance of these yeast findings, we evaluated C1311 singly and in doublet combination with paclitaxel in human bladder cancer in the *in vivo* hollow fiber assay and observed efficacy. By applying COXEN to gene expression data from 40 bladder cancer cell lines and 30 human tumors with associated clinical response data to platinum-based chemotherapy, we provide evidence that signatures of C1311 sensitivity exist within nonresponders to this regimen. Coupling COXEN and yeast chemigenomics provides rational combinations with C1311 and tumor genomic signatures that can be used to select bladder cancer patients for clinical trials with this agent.

Neoplasia (2011) 13, 72–80

Introduction

Bladder cancer is common and costly [1]. Nearly 30% of patients present with muscle invasive bladder tumors at diagnosis, and approximately 50% of these patients develop distant recurrence and require systemic chemotherapy [2]. With standard platinum combination therapy (commonly cisplatin or carboplatin and gemcitabine, GC), a median survival of only 13 months can be achieved in patients with advanced disease, with modest response rates reported for second line agents for treatment failures [3].

We have recently reported an informatics approach termed *COXEN*, for coexpression extrapolation, that uses cell line transcriptional signatures and associated *in vitro* sensitivity to therapeutic compounds to

Address all correspondence to: Dan Theodorescu, MD, PhD, University of Colorado Comprehensive Cancer Center, 13001 E 17th Pl MS #F-434, Aurora, CO 80045. E-mail: dan.theodorescu@ucdenver.edu

[1]This work was supported by the National Institutes of Health grant CA075115. The National Institutes of Health had no role in study design, data collection, and analysis, decision to publish, or preparation of the article.

[2]This article refers to supplementary materials, which are designated by Figure W1 and Tables W1 to W4 and are available online at www.neoplasia.com.

Received 22 August 2010; Revised 7 October 2010; Accepted 9 October 2010

Copyright © 2011 Neoplasia Press, Inc. All rights reserved 1522-8002/11/$25.00
DOI 10.1593/neo.101214

predict sensitivity of independent cell line panels and patient responses to such agents [4]. The novel aspect of this approach is its ability to select from sensitivity biomarker genes derived from cell lines a subset that maintain concordant expression in a second cohort of cell lines or human tumor samples. Importantly, this analysis is done *a priori*, without knowledge of the pattern of sensitivity or clinical response is in the second set. Originally reported as predictive of the outcomes of separate clinical studies in 84 patients [5], recently this algorithm has been used successfully to stratify clinical outcomes nearly 500 patients with diverse tumor types [6].

One COXEN application we have reported is in drug discovery, using publicly available data for 45,545 compounds from the US NCI Developmental Therapeutics Program's screen of 60 cell lines from nine different tumor histologies (NCI-60) [7]. Because bladder cancer cell lines were not part of the NCI-60, we used COXEN to predict which of the 45,000 drugs would be highly active in human bladder cancer. Top hits from this analysis included NSC-637993 and NSC-645809 (C1311), two imidazoacridinone class compounds. Evaluation of NSC-637993 on a panel of 40 bladder cancer cell lines indicated that more than 60% exhibited 50% growth inhibition at the micromolar level or better [5].

The imidazoacridinones are a promising new class of compounds for human cancer [8] and are thought to function through several mechanisms. Work with C1311 suggests that its mechanisms of action may include DNA intercalation [9], as well as inhibition of topoisomerase II [10] and the FLT3 tyrosine kinase [11]. Recent studies additionally suggest that the mechanism of inhibition of topoisomerase II may be due to C1311 interfering with ATP binding to the enzyme, perhaps in a fashion analogous to its inhibition of FLT3 [12]. Given our *in vitro* results with NSC-637993 in bladder cancer cells and promising results obtained for C1311 in early clinical trials in other tumor types [13,14], we decided to perform a preclinical evaluation for these two related molecules in bladder cancer with the intent to pave the way for future clinical trials with these agents.

Because essentially no cures are observed in the setting of second-line therapy for metastatic disease treated with single agents [3], we applied yeast chemical genetics methods to define and then validate in human bladder cancer, rational combination therapy with C1311. In addition, given our success with COXEN-based gene expression signatures in predicting chemotherapeutic outcomes, we also provide evidence that, among patients who fail first-line platinum chemotherapy for metastatic bladder cancer, there exists a cohort that exhibits transcriptional signatures suggestive of response to C1311.

Materials and Methods

Cell Culture, In Vitro*, and* In Vivo *Drug Sensitivity*

All human bladder cancer cells (BLA-40 panel), culture conditions, and our protocol for assay of drug sensitivity have been reported previously [5,15]. IC_{50} values (concentrations capable of inducing 50% inhibition of cellular growth) were calculated for the 40 cell lines using an improved Spline-fitting approach in the statistics suite, R (www.R-project.org). *In vivo* sensitivity studies used the hollow fiber assay (HFA), reported before [16], and in Supplementary Methods. The significance of growth inhibition in HFA results was tested by single-sample *t* tests against the hypothesis that there was no inhibition, in PRISM (GraphPad Software, La Jolla, CA), with two-sided *P* values reported.

Competitive Yeast Growth Experiments

In *Saccharomyces cerevisiae*, mutant strains with knockouts of all non-lethal open reading frames (ORFs; ~4600) are available that are tagged with two unique oligonucleotide "barcodes" that are flagged by universal polymerase chain reaction primers for detection through microarrays, as detailed before [17]. For competitive growth experiments, the collection of homozygous diploid mutant cells (EUROSCARF; Institute of Molecular Biosciences, Frankfurt, Germany) were grown on YPD agar containing G418, pooled and frozen in 0.23-ml aliquots at OD = 21.5. For YPD growth, cells were diluted to 6.17 × 10 E7 cells/ml and grown to saturation (five generations). Cultures were sequentially diluted to 6.17 × 10 E7 cells/ml for consecutive growth experiments (10, 15, and 20 generations). C1311 stocks were maintained at 100 μM in DMSO, and cells were treated with 0, 1, or 5 μl in YPD plus 1% DMSO. Benomyl treatment was at 15 μg/ml in YPD plus 1% DMSO. Genomic DNA was recovered using MasterPure yeast DNA purification kit (EPICENTRE Biotechnologies, Madison, WI) and hybridized to Affymetrix Yeast TAG4.0 microarrays (Affymetrix Inc, Santa Clara, CA) as per the manufacturer's instructions.

Analysis of Yeast TAG Array and Synthetic Lethal Data

The Affymetrix Yeast TAG4.0 array data were analyzed using the software developed by the Giaever laboratory, which normalizes, quality filters, and background adjusts data as detailed in the Supplementary Methods and previous publication [18]. The Yeast TAG4.0 drug data as well as synthetic lethal data were binarized, assigning a 1 to synthetic lethal query-target pairs and 0 to all other ORF pairs. Combining the drug and synthetic lethal data resulted in a binary matrix with 1521 rows of yeast query genes and 6 drug treatments (4 C1311 and 2 benomyl) and 2804 columns of yeast target genes. We note that the original size of each drug binary vector was 6431 (i.e., the number of yeast deletion strains interrogated on the array) and reduced to 2804 after being projected onto the set of available yeast target genes. These data were clustered in two dimensions (i.e., cluster both rows and columns) with a cosine distance metric to this 1521 × 2804 binary matrix using the *clustergram* function in MATLAB Version 7.9.0 (MathWorks, Natick, MA). Lists of yeast strains with reduced fitness for benomyl and C1311 were examined for statistically significant enrichment of gene ontology terms by GO::TermFinder [19] using default settings.

Drug Sensitivity Correlation Analyses

We calculated Spearman correlations between C1311 and NSC-637993 compound data, paired distributions for C1311 and NSC-637993 were compared using the Wilcoxon matched-pairs test, in MATLAB and PRISM, respectively. For correlation analysis of the ~4600 developmental therapeutics program drugs [20] to C1311 across the NCI-60, we calculated Spearman correlation coefficients of cell line IC_{50} values for C1311 to all other drugs, calculated Benjamini-Hochberg–corrected *P* values, and used Kernel-Smoothing function to plot the distribution of correlation coefficients, all in MATLAB.

Training and Testing COXEN-Based Classifier

Gene expression profiling of the BLA-40 bladder cancer panel [GEO:GSE5845] [15] and that of the NCI-60 panel [GEO: GSE5720] [21] and *in vitro* testing data for the compounds [22] were used. For detailed methods on interplatform prediction of C1311 sensitivity and biomarker selection, see Supplementary Methods.

Results

Cytotoxicity of Imidazoacridinones on Human Bladder Cancer Cells In Vitro *and* In Vivo

Having discovered that NSC-637993 exhibits activity in many bladder cancer cell lines [5], we were interested whether the related imidazoacridinone, C1311, might also have activity in bladder cancer. We generated dose-response curves for C1311 in our 40 bladder cancer cell lines (BLA-40) [15] across a concentration range of five logs, estimated IC_{50} concentrations, and compared them to IC_{50} values for NSC-637993 for the same panel. We observed robust activity of C1311 in these cells, with IC_{50} values uncorrelated to the expression of TOP2A ($r_s = -0.11$, $P = .52$) and FLT3 ($r_s = -0.26$, $P = .11$), putative targets of C1311 (Figure 1*A*; for complete data, see Table W1 and Figure W1).

Given the structural and functional similarity of C1311 to NSC-637993, we wished to determine whether IC_{50} values for NSC-637993 and C1311 were correlated. We found that the IC_{50} values of the two drugs were significantly correlated (Figure 1*B*, $r_s = 0.44$, $P = .006$), whereas there was no significant difference between the paired distributions of IC_{50} values for C1311 and NSC-637993 ($P = .42$) across these cells. This is consistent with data for the NCI-60 panel of cells, for which a similar correlation was also observed ($r_s = 0.51$, $P < .01$), as well as a trend toward superiority of C1311 over NSC-637993 ($P < .01$; not shown). In addition, we observed that C1311 compares favorably with cisplatin and gemcitabine, the standard-of-care drugs for bladder cancer (Figure 1*C*).

We next tested these compounds *in vivo* using HFA [16]. On the basis of our *in vitro* characterization of sensitivity to C1311 and NSC-637993 (Table W1) as well as evaluation of cell lines for compatibility with the HFA, we selected one cell line exhibiting a low IC_{50} (T24T, sensitive), two cell lines with intermediate IC_{50} values (FL3 and UMUC1), and one cell line with a high IC_{50} (KK47, resistant) to C1311, as indicated in Table 1 for HFA studies. Mice were treated four times daily at 20 mg/kg by intraperitoneal injection, with animals killed at 96 hours for quantitative measurement of cell growth. Compared with untreated animals, we found that, in three of four cell lines tested, there was a significant inhibition of growth (all $P \leq 0.02$), the exception being the resistant KK47 cell line (Table 1). Similar results were shown testing NSC-637993 in this assay (Table W2*A*).

Chemigenomic Profiling in Yeast Suggests a Mode of Action for C1311

Given the wide variety of potential targets reported for C1311 and the imidazoacridinone class [8] as well as the need to better characterize C1311's mechanism for potential rational drug combinations [23], we embarked on an unbiased, new strategy using budding yeast to begin to characterize C1311's mechanism. Recent advances in yeast genetics enable high-throughput screening of yeast ORF deletion strains for those that are sensitized to compounds or for synthetic lethal relationships between two deletion mutants [24]. Formally, the analyses are comparable; a compound may be effective because it inactivates a gene product and is therefore similar to deletion of the ORF. Such analyses have even yielded promising results for inference of conserved cellular pathways perturbed by drugs by comparing the pattern of strains sensitized to a drug to genome-wide genetic synthetic lethal relationships [25].

We grew the pooled yeast deletion mutants in the presence of two increasing concentrations of C1311 and profiled the strains remaining after 10 and 20 generations of growth, compared with control-treated pools, using Yeast TAG4.0 microarrays. For the lower concentration of C1311, we found that 27 and 32 yeast strains showed significantly reduced fitness when assayed after 10 and 20 generations. As expected, at higher a concentration, we found more strains displaying reduced

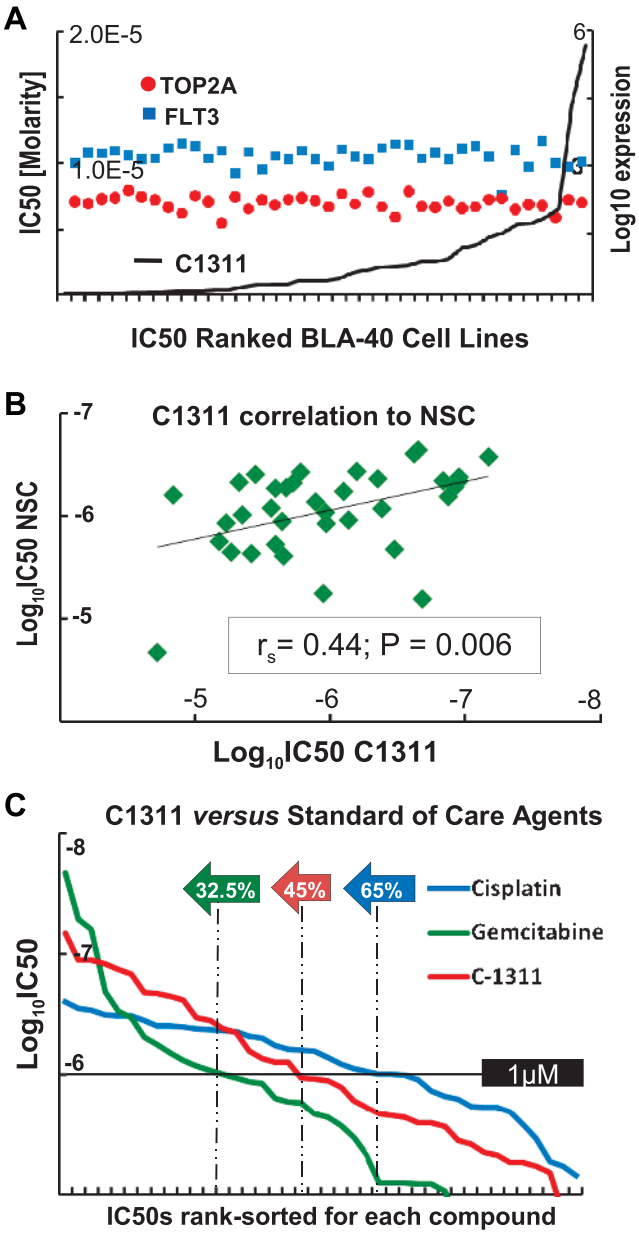


Figure 1. C1311 and NSC exhibit similar, favorable *in vitro* activities, comparable to standard-of-care agents. (A) IC_{50} values for C1311 were determined by Spline regression for the BLA-40 cell line panel, plotted ranked ordered from left to right, then each corresponding cell line's expression of TOP2A and FLT3 expression (Affymetrix probes 201291_s_at and 206674_at, both $\log_{10}$ values for visualization in scale). (B) Scatter plot of NSC compound (ordinate) and C1311 (abscissa) IC_{50} values across the BLA-40 panel, nonparametric Spearman correlation, and *P* value. (C) Comparison of C1311 and standard-of-care drugs by IC_{50} across the BLA-40 panel. IC_{50} values C1311, cisplatin, and gemcitabine were rank ordered for the 40 cell lines for each drug and plotted in ascending order on the log-scale *y* axis. The green, pink, and blue arrows indicate the percentage of the BLA-40 cell lines that exhibit IC_{50} values below the 1-μM range, demonstrating that C1311 exhibits similar activity to agents currently in clinical use.

Table 1. HFA Results for C1311.

Cell Line	C1311 Status*	Log_{10} IC_{50}	SQ†	P‡	IP§	P‡	Overall	Overall P‡
T24T	Sensitive	−6.62	65.5	<.0001	71.0	.0006	68.3	<.0001
FL3	Intermediate	−5.45	78.6	.014	58.3	<.0001	68.4	.0002
UMUC1	Intermediate	−5.35	97.7	.067	60.5	<.0001	79.1	.02
KK47	Resistant	−4.84	80.1	<.0001	101.8	.57	90.9	.10

*Relative sensitivity to C1311 of indicated cell line. Of cell lines adaptable to the hollow fiber assay, four cell types of varying *in vitro* sensitivities were selected for validation *in vivo*.
†Average percentage of control growth across four replicates at the subcutaneous implantation site.
‡Two-tailed P value for single-sample t test against the hypothesis that the inhibition was 0%.
§Average percentage of control growth across four replicates at the intraperitoneal implantation site.

fitness, 32 and 49, after 10 and 20 generations of growth, respectively. We also found 12 (62.5-fold over mean expected by random chance) and 15 (42.5-fold over mean expected by random chance) strains in common when comparing the low- and high-concentration data at 10 and 20 generations, respectively, both $P \ll .0001$. The highly nonrandom concordance in reduced fitness strains identified by separate growth replicates treated at two different drug concentrations illustrates the significant reproducibility of the results. The nonredundant union of the strains that displayed reduced fitness at both concentrations and generations yielded 91 strains (Table W3*A*).

Combining our C1311 data with the 25,540 synthetic lethal interactions identified in yeast (as of November 24, 2009 [17]) resulted in a binary matrix with 1521 rows of yeast query genes. To cluster these data in an interpretable way and allow comparison of the pattern of strains sensitized to C1311 to genome-wide synthetic lethal interactions, as has identified pathways targeted by drugs before [25], we used two-dimensional hierarchical clustering, as shown in Figure 2, *A* to *C*. The C1311 replicates (two concentrations after 10 or 20 generations of competitive growth) cluster immediately next to each other into a region of the *clustergram* enriched with ORFs involved with membrane lipid biogenesis (*erg2 erg6 erg24* and *erg28*; Figure 2*B*). We interpreted this as suggesting a function for C1311 in perturbing cellular lipid biosynthesis or membrane function. Also consistent with this finding, using gene ontology to evaluate the 91 C1311 reduced fitness strains (Table W3*A*), we found a highly significant enrichment of gene ontology terms related to several lipid biogenesis pathways (Table 2*A*).

Chemigenomic Profiling in Yeast Identifies Known Taxane Targets

New anticancer agents are tried first in the setting of primary treatment failure, meaning a potential future trial would use C1311 alone or a C1311-based combination regime after failure of GC [26]. Paclitaxel, a taxane, has shown substantial activity both alone and in combination therapies for bladder cancer [27] and is the leading second-line agent in practice today. Hence, a doublet combination of C1311 with paclitaxel would seem appealing.

Evaluation of paclitaxel in the chemigenomic assay would provide two significant insights. First, it would validate our chemigenomic approach because the molecular target of taxanes is known, supporting the premise that deletion strains sensitized to C1311 (Table W3*A*) identify its true mode of action. Second, if comparison of such unbiased evaluation of taxanes mode of action with that of C1311 reveals little overlap, such information would support combination treatment with these two drugs [23]. Here we use benomyl, an antitubulin drug that inhibits β-tubulin–like paclitaxel because the latter does not bind yeast β-tubulin because of the slight differences in the sequences of the proteins between yeast and humans [28].

Using a sublethal concentration of benomyl to treat pooled yeast deletion mutants and assaying by the same microarray platform after 10 or 20 generations of competitive growth, we found 32 and 16 significantly reduced strains (a union resulting in 34 strains; Table W3*B*). These data cluster immediately beside each other in a region of the *clustergram* that is highly enriched in genes that function in the mitotic spindle and immediately beside the *tub3* tubulin deletion mutant (Figure 2*C*). These findings provide "proof-of-principle" because the pattern of synthetic sensitivities caused by inhibiting microtubule function with benomyl is most similar to inactivating tubulin with through deletion mutation, confirming that our approach may define modes of action of drugs. Also supporting this finding, when we used GO::TermFinder, the 34 benomyl-sensitized strains showed significant enrichment of terms relating to tubulin complex formation, among others (Table 2*B*).

Nonoverlapping Pathways and Sensitivities for C1311 and the Taxane Benomyl

To examine whether the strains identified as synthetic lethal to C1311 and benomyl (Table W3) significantly overlapped, we used the χ^2 test, finding no significant overlap (two strains, $P = .14$, Yates corrected). Given that the strains that showed reduced fitness to the drugs were essentially mutually exclusive, these findings are consistent with C1311 and benomyl having distinct modes of action in yeast. However, to provide additional support for the strategy of combining imidazoacridinones with taxanes, we availed ourselves to published data encompassing ~4400 drugs tested across the aforementioned NCI-60 cell line panel [20]. Such data afford the opportunity to test correlation of each drug's pattern of IC_{50} values across the 60 cell lines to that of C1311, allowing objective comparisons of patterns of each drug to C1311 but also relative comparisons among a large number of diverse bioactive compounds. Consistent with their targeting disparate cellular pathways, we found that paclitaxel was nonsignificantly correlated to C1311 at a level of $r_s = 0.25$ ($P = .11$), essentially indistinguishable from the average correlation of $r_s = 0.24$ ($P = .27$) across all drugs. Interestingly, even the standard-of-care, GC doublet drugs, cisplatin, and gemcitabine were more significantly correlated to C1311 than paclitaxel ($r_s = 0.56$, $P < .0001$, $r_s = 0.58$, $P < .0001$, respectively). Figure 2*D* shows these findings graphically, plotted against the ranked distribution of correlations of the ~4400 drugs. Taken together, these findings from yeast and human cells suggested preclinical evaluation of a paclitaxel C1311 doublet therapy, which we undertook below.

In Vivo *Evaluation of Imidazoacridinone-Taxane Combination Therapy in Human Bladder Cancer*

For C1311 or NSC-637993 plus paclitaxel doublet HFA experiments, we used our previously reported *in vitro* IC_{50} data for paclitaxel in the BLA-40 panel [15] to select cell lines evincing several informative combinations of paclitaxel *versus* C1311 sensitivities as indicated in Table 3. We used a cell line with low IC_{50} values to both the imidazoacridinones and paclitaxel (UMUC6), two cell lines with intermediate IC_{50} values to both (T24 and 5637), and KK47 that had high IC_{50} values to both drug classes. Animals were treated by intraperitoneal injection four times with C1311 or NSC (20 mg/kg) plus paclitaxel (15 mg/kg) and sacrificed at 96 hours. Comparing untreated and doubly treated animals, we found that combinations were effective against all cell lines (all $P < .01$), including KK47, which was resistant to imidazoacridinone monotherapy (Table 3), with similar findings for

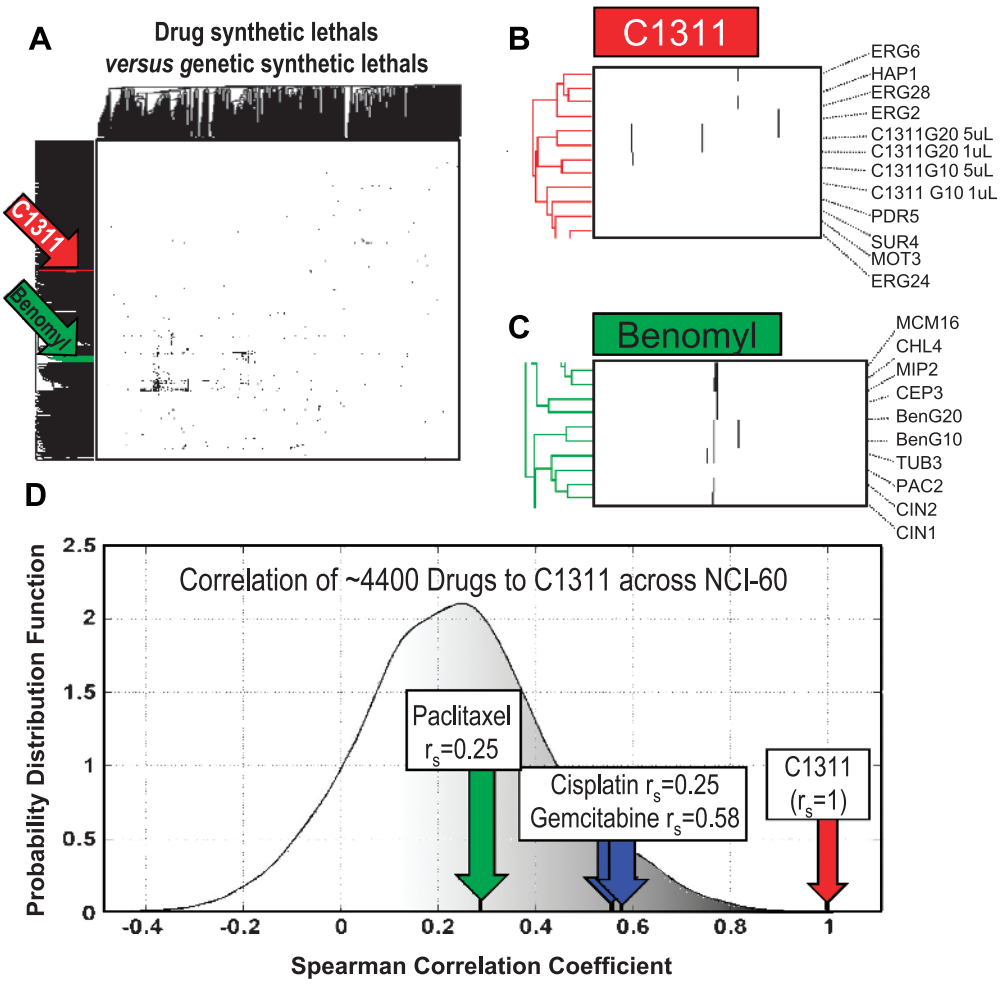


Figure 2. Chemigenomic analysis of C1311 and paclitaxel between yeast and human cells. (A) Two-way hierarchical cluster of a 1521 × 2804 binary matrix where a black pixel represents either a synthetic lethal interaction between two yeast ORFs or reduced fitness between a drug and a yeast deletion mutant, or a white pixel represents all other two-ORF or drug-ORF pairs. We highlight the C1311 (red) and benomyl (green) clusters. (B) We see that C1311 is in a relatively sparse area of the matrix, and the four treatments with C1311 (two concentrations after 10 or 20 generations of competitive growth) cluster immediately next to each other and among several deletion strains involved with membrane lipid biogenesis, implicating this pathway in the function of C1311. (C) Enlarged inset view of the benomyl cluster from (A) showing benomyl treatments for 10 and 20 generations and neighboring deletion strains, enriched for microtubule and spindle components, as would be expected for this microtubule poison. (D) IC_{50} patterns across 60 cell lines for ~4400 drugs from the 60 cell lines of NCI-60 screen [20] were correlated to that of C1311 and the probability distribution function of the coefficients was plotted. Paclitaxel, cisplatin, and gemcitabine exhibited the indicated correlation coefficients.

combinations with NSC-637993 and paclitaxel (Table W2*B*). These findings suggest that despite the significant differences between *in vitro* screening and quantitative evaluation of the pharmacologic capacity of the drugs to reach subcutaneous and intraperitoneal compartments *in vivo*, C1311 (and NSC) remain highly effective, whereas addition of paclitaxel expands the spectrum of cell lines inhibited significantly by treatment.

Development of a Predictive Biomarker Model of C1311 Sensitivity

Taken together, the promising *in vitro* and *in vivo* results for C1311 suggest the possibility of testing it in a future clinical trial, a setting where significant efficiencies may achieved by biomarker-guarded selection of patients most likely to respond to therapy [29]. In particular, if gene expression signatures suggestive of sensitivity to C1311 were present in patients who did not respond to cisplatin-based therapy, such data would further support its evaluation. We examined this question using the COXEN algorithm, which develops, based on drug sensitivity and gene expression profiling in cell lines, gene expression model (GEM) predictors of drug response in patients [4].

To derive a gene expression signature of C1311 sensitivity, we began by selecting candidate sensitivity biomarkers by rank-based correlation (r_s = 0.4) of C1311 IC_{50} values across the NCI-60 panel, finding that 219/22283 Affymetrix HG-U133A microarray probes meet this criterion (false discovery rate = 0.1 by random permutation testing, for probe information and correlation coefficients, see Table W4). Evaluation of the potential functional associations of these 219 probe sets using the Ingenuity Pathway Analysis program identified the glycerophospholipid metabolism pathway as the most significantly enriched, supporting in cancer cells our observations associating C1311 with lipid biogenesis in budding yeast. Linoleic and arachidonic acid metabolic pathways were also high scoring pathways (Table 4).

Table 2. Gene Ontology Term Enrichment among C1311- (A) and Benomyl- (B) Sensitized Yeast Strains.

GOID*	GO_Term*	Cluster Frequency	Background Frequency	$P^{\dagger}$
(A)				
6629	Lipid metabolic process	19/90 genes, 21.1%	259/7166 background, 3.6%	1.17E–07
44255	Cellular lipid metabolic process	18/90 genes, 20.0%	254/7166 background, 3.5%	6.51E–07
8610	Lipid biosynthetic process	12/90 genes, 13.3%	140/7166 background, 2.0%	5.66E–05
9987	Cellular process	81/90 genes, 90.0%	4819/7166 background, 67.2%	.00015
16126	Sterol biosynthetic process	6/90 genes, 6.7%	33/7166 background, 0.5%	.00111
6694	Steroid biosynthetic process	6/90 genes, 6.7%	33/7166 background, 0.5%	.00111
9058	Biosynthetic process	46/90 genes, 51.1%	2011/7166 background, 28.1%	.00112
44249	Cellular biosynthetic process	45/90 genes, 50.0%	1994/7166 background, 27.8%	.00239
16129	Phytosteroid biosynthetic process	5/90 genes, 5.6%	26/7166 background, 0.4%	.0059
6696	Ergosterol biosynthetic process	5/90 genes, 5.6%	26/7166 background, 0.4%	.0059
16125	Sterol metabolic process	6/90 genes, 6.7%	46/7166 background, 0.6%	.0082
6631	Fatty acid metabolic process	6/90 genes, 6.7%	46/7166 background, 0.6%	.0082
8202	Steroid metabolic process	6/90 genes, 6.7%	46/7166 background, 0.6%	.0082
16128	Phytosteroid metabolic process	5/90 genes, 5.6%	28/7166 background, 0.4%	.00864
8204	Ergosterol metabolic process	5/90 genes, 5.6%	28/7166 background, 0.4%	.00864
(B)				
43486	Histone exchange	5/33 genes, 15.2%	10/7166 genes, 0.1%	4.52E–08
7021	Tubulin complex assembly	4/33 genes, 12.1%	5/7166 genes, 0.1%	2.24E–07
43044	ATP-dependent chromatin remodeling	5/33 genes, 15.2%	29/7166 genes, 0.4%	2.00E–05
7023	Post-chaperonin tubulin folding pathway	3/33 genes, 9.1%	5/7166 genes, 0.1%	.00011
34728	Nucleosome organization	5/33 genes, 15.2%	51/7166 genes, 0.7%	.00037
43933	Macromolecular complex subunit organization	11/33 genes, 33.3%	470/7166 genes, 6.6%	.00054
34621	Cellular macromolecular complex subunit organization	10/33 genes, 30.3%	392/7166 genes, 5.5%	.00077
6338	Chromatin remodeling	5/33 genes, 15.2%	76/7166 genes, 1.1%	.00268
6457	Protein folding	5/33 genes, 15.2%	86/7166 genes, 1.2%	.00488

*GOID and GO Terms from the Gene Ontology Consortium, www.geneontology.org, identified by GO::TermFinder.
†Bonneferoni-corrected *P* value for enrichment of indicated term among yeast ORF knockout strains identified as exhibiting reduced fitness on growth with C1311 or benomyl (cluster frequency) to the background frequency of such GO terms in the genome. All GOIDs presented were associated with an FDR approximating 0%.

We then applied a critical aspect of the COXEN algorithm to uncover which of the 219 probe sets were concordantly expressed across three data sets (NCI-60, BLA-40, and a published bladder tumor data set [30]) and then derive a GEM predicting C1311 sensitivity from them. The human tumor data set was included so that the model could be applied on human tumor data sets, as we have reported before [6]. We systematically examined subsets of the 219 C1311-associated probes that maintained concordant expression between the three data sets as described in Supplementary Methods 1, for performance in predicting sensitivity of BLA-40 cells based on the similarity of their gene expression to the NCI-60 cells used for training. The maximally performing subset of five probes (Table W4) exhibited highly concordant expression between all three of the aforementioned data sets and was implemented in a weighted *k* nearest-neighbor (weighted kNN) classifier in our final GEM [31]. This strategy assigned a prediction for each BLA-40 test sample, based on the correlation of it its expression of the five probes, to sensitive and resistant groups of NCI-60 cells. The GEM exhibited significant ability to predict the sensitivity of the BLA-40 from the NCI-60 ($P = .01$; Figure 3*A*). For full details on the development of GEM, please see Supplementary Methods 1.

Evaluation of the C1311 GEM in Patients Undergoing Platinum-Based Chemotherapy

To test whether the GEM-identified signatures of sensitivity among patients showed resistance to cisplatin-based therapy, we tested it on the microarray data for a cohort of 30 patients, published by Als et al. [32], where the response to standard cisplatin-based chemotherapy was known. We also used it to evaluate another reported microarray study by Sanchez-Carbayo et al. [30] to examine the association of the GEM's predictions with other clinicopathologic characteristics. Both studies profiled histologically verified, fresh-frozen primary tumor tissues (biopsies and surgical resection specimens, respectively) on the Affymetrix HG-U133A platform. Figure 3*B* demonstrates through hierarchical clustering how the two cell lines and two human tumor data sets are capable of being clustered together in an interpretable fashion across the five COXEN-selected C1311 sensitivity GEM genes.

Table 3. HFA Results for C1311 and Paclitaxel.

Tx	Cell Line	C1311 Status*	Log_{10} IC_{50}	Paclitaxel Status*	Log_{10} IC_{50}	SQ†	$P^{\ddagger}$	IP§	$P^{\ddagger}$	Overall	Overall $P^{\ddagger}$
C1311 + Paclitaxel	UMUC6	Sensitive	−6.92	Sensitive	−8.99	66.1	.0003	51.2	<.0001	58.7	<.0001
	HTB9	Intermediate	−5.64	Sensitive	−8.93	40.5	.065	74.0	.0024	57.2	.0047
	T24	Intermediate	−5.66	Intermediate	−7.36	76.9	.002	55.7	<.0001	66.3	<.0001
	KK47	Resistant	−4.84	Resistant	>−7	70.9	<.0001	59.4	.0002	65.1	<.0001

*Relative responsiveness to C1311 or paclitaxel of indicated cell line. Of cell lines adaptable to the hollow fiber assay, four cell types of varying *in vitro* combinations of sensitivities to both drugs were selected for validation *in vivo*. We have reported *in vitro* sensitivities to paclitaxel across the BLA-40 panel before [15].
†Average percentage of control growth across four replicates at the subcutaneous implantation site.
‡Two-tailed *P* value for single-sample *t* test against the hypothesis that the inhibition was 0%.
§Average percentage of control growth across four replicates at the intraperitoneal implantation site.

Table 4. Canonical Pathways Enriched in C1311-Correlated Microarray Probes.

Ingenuity Canonical Pathways*	$P^{\dagger}$	Ratio	Molecules
Glycerophospholipid metabolism	.004	3.11E–02	PPAP2B, PLA2R1, CHKA, AGPAT1, PGS1, PCYT1A
N-glycan biosynthesis	.028	3.23E–02	B4GALT4, ST6GAL1, MAN1C1
Riboflavin metabolism	.028	3.64E–02	ACP5, ACP1
Arachidonic acid metabolism	.045	1.76E–02	CYP2F1, CYP1A1, PLA2R1, PLOD1
Linoleic acid metabolism	.048	2.42E–02	CYP2F1, CYP1A1, PLA2R1
Sphingolipid metabolism	.059	2.68E–02	NAGA, PPAP2B, ASAH1
AMPK signaling	.100	2.41E–02	MTOR, ACACB, STRADA, PPM1D
Glycerolipid metabolism	.104	1.92E–02	NAGA, PPAP2B, AGPAT1
EIF2 signaling	.107	3E–02	EIF2C2, EIF2AK2, GSK3B

*Ingenuity Pathway Analysis (IPA), Version 8, www.ingenuity.com.
†*P* values from IPA are Benjamini-Hochberg corrected.

Using the weighted kNN classifier to classify the human tumors, we found in the chemotherapy study of Als et al. that distributions of our C1311 prediction scores showed no differences among the study's observed chemotherapy response groups ($P = .62$; Figure 3*C*), suggesting that predictions were not reflective of general drug resistance and supportive of the idea that candidates for C1311 treatment exist in nonresponders to standard-of-care drugs. In addition, in both studies of Sanchez-Carbayo et al. and Als et al., we observed no difference in C1311 sensitivity prediction scores based on patient age, sex, tumor grade, or stage (data not shown). Furthermore, we did not observe significant associations between our C1311 predictions and survival in either data set ($P = .46$ and $P = .54$, respectively; Figure 3*D*). These findings again suggest that the C1311 predictions are also not associated with general phenotypes, like aggressiveness of tumors, and are independent of traditional pathologic factors and outcome, an important requirement for molecular assays to have clinical utility [6].

Discussion

Efficient drug development relies on the identification of candidate compounds, their preclinical validation in model systems, and translation in clinical trials. The availability of data from high-throughput technologies such as drug screens on cell panels [7] or publicly available gene expression profiling [33] provides building blocks that can be synthesized with informatic tools such as COXEN to provide an

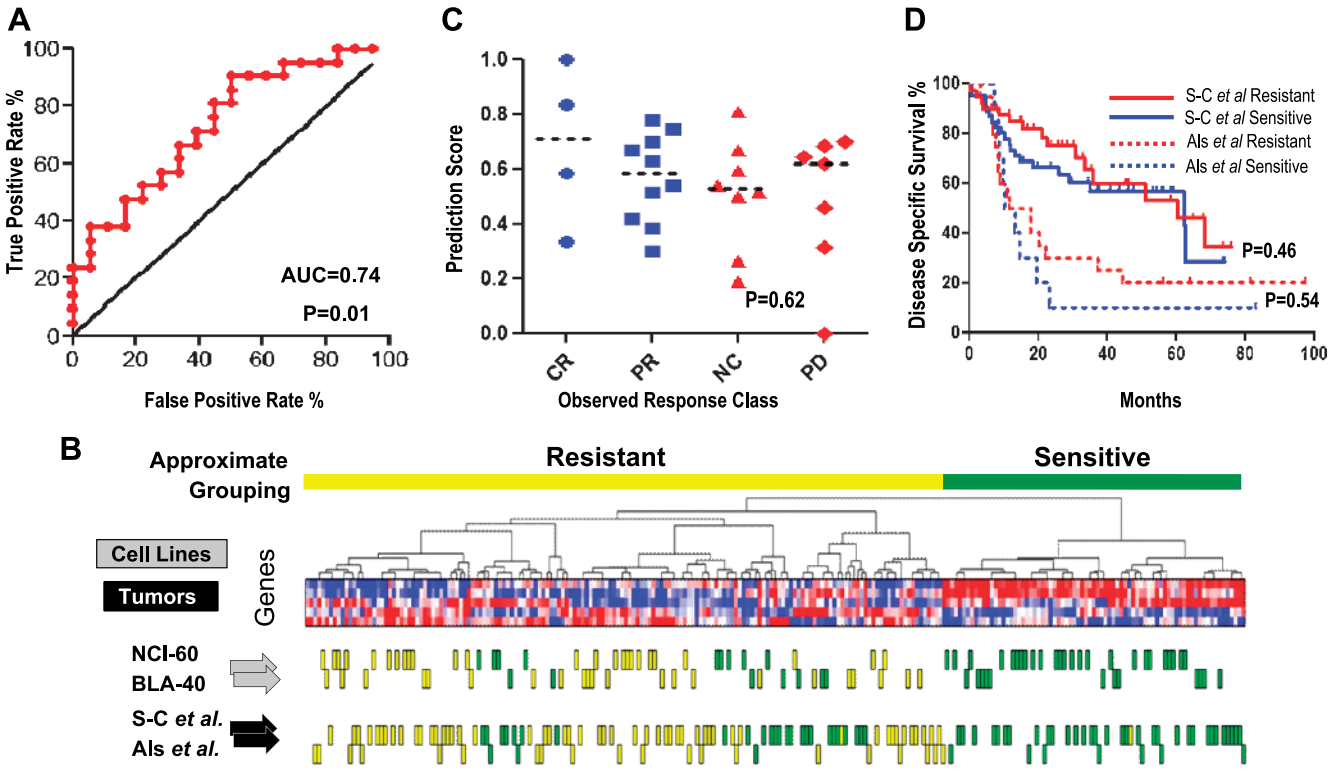


Figure 3. Prediction of C1311 sensitivity between cell line panels and in human tumors. (A) COXEN analysis was used to develop a set of probe sets associated with sensitivity to C1311 in the NCI-60 cell line panel and concordantly expressed between the NCI-60 panel, the BLA-40 panel, and Sanchez-Carbayo et al. tumor gene expression data sets. Then, a nearest neighbor–based classification approach used to classify the BLA-40 cell line panel based on the NCI-60 panel, and the ROC curve was plotted for the classes assigned (sensitive or resistant) to test its ability to discriminate (area under the curve = 0.74, 95% CI = 0.59-0.90, $P = .01$). (B) Clustering of multiple data sets by C1311 sensitivity genes. A two-dimensional hierarchical cluster of NCI-60, BLA-40, Sanchez-Carbayo et al., and Als et al. data sets across 28 three-way concordant probe sets. Individual NCI-60 (actual) and BLA-40 (predicted) cells are indicated by boxes, showing resistance in yellow and sensitivity in green. Individual Sanchez-Carbayo et al. and Als et al. (both also predicted) tumor data sets are also indicated as for the BLA-40 cells. The *clustergram* illustrates how the COXEN methodology may select concordant biomarkers between platforms such that such gene expression patterns allow visualization or computational prediction of interpretable relationships between diverse biologic systems. (C) C1311 prediction values were dot-scatter plotted for response classes from the Als et al. data set, including CR (complete responder), PR (partial responder), NC (no change), and PD (progressive disease), finding no significant difference by nonparametric analysis of variance. (D) Kaplan-Meier analysis of survival by C1311 prediction indicates no systematic association between C1311 prediction class and survival in either study.

integrated pipeline for drug development, an example of which we report herein. We previously reported the discovery of NSC-637993 as a promising candidate compound for bladder cancer through the COXEN algorithm; the first imidazoacridinone class drug to be so evaluated for bladder cancer [5]. Among the top hits was a related imidazoacridinone, C1311, with favorable activity, toxicity, and tolerability profiles [13,14,34]. Not surprisingly, we observed a correlation of C1311 to NSC-637993 responsiveness, as well as a similar range of IC_{50} values on bladder cell lines. Interestingly, our chemigenomic screen using budding yeast suggested that C1311's mechanism may involve lipid biosynthesis pathways, a novel observation adding to the large number of potential targets postulated for the drug in recent reports [9,11,35–37]. This unexpected finding was nonetheless supported in mammalian cells by our observed enrichment of probes for genes involved in glycerophospholipid metabolism among those correlated to C1311 IC_{50} values across the NCI-60 cell line panel. Our group is currently using these data to attempt to identify the target and sensitizing agents for C1311.

The observation of essentially mutually exclusive patterns of synthetic lethality between benomyl and C1311 provided a testable therapeutic combination because of the similarity between benomyl and the approved chemotherapeutic agent, paclitaxel [28]. Paclitaxel is especially useful because it has complementary toxicities to those of C1311 and has been used as monotherapy [15] in patients with advanced bladder cancer that have failed platinum agents [3]. Supporting our observations of different classes of yeast knockouts conferring sensitivity to the two drugs, correlation of sensitivity patterns of C1311 across the NCI-60 cell lines to 4463 drugs [20] found that C1311 and paclitaxel were not correlated and even less correlated than C1311 to cisplatin or gemcitabine. Also supportive of this concept is a previous report of activity of C1311 in advanced breast cancer failures that included taxane failures [37]. We performed an *in vivo* evaluation for both C1311 and NSC-637993 using the National Institutes of Health/National Cancer Institute–developed HFA and found that both agents were effective alone in most bladder cancer cells. Importantly, combining these with paclitaxel in cells found to be resistant to imidazoacridinone monotherapy, such as KK47, led to significant inhibition. Taken together, these data suggest that clinical evaluation of C1311 with or without paclitaxel in the setting of cisplatin-based treatment failures is warranted.

To optimize patient selection for bladder cancer clinical trials with C1311, a biomarker for sensitivity to this drug is needed. In particular, retrospective examination of such a biomarker or prediction model on gene expression data from patients who have already been treated with platinum regimens would provide an indication whether there is cross resistance to C1311. Because COXEN-based classifiers have been shown to be predictive of outcome in nearly 500 patients from nine clinical trials [5,6], we used this technology to develop a GEM predicting response to C1311. We found that COXEN predictions did not differ significantly based on patient age, sex, tumor grade, or stage in two different data sets of patients [30,32]. Making predictions on the study of Als et al., which includes standard cisplatin-based therapy response and survival outcome data, we found that predictions did not differ significantly between patients evincing complete response, partial response, no change, or progressive disease. An important limitation of these findings is that, without data from an actual trial, it is not possible to assign a cutoff value for prediction scores that definitively identifies a responder or nonresponder to C1311. However, these findings do suggest that our C1311 predictions were not simply an index of tumor aggressivity or general drug resistance. Taken together, they provide the rational framework for developing a future (biomarker-selected or correlated) clinical trial of C1311 in the clinical setting of cisplatin-based treatment failures.

In summary, we demonstrate that combining COXEN and yeast chemigenomics allows formulation of rational drug combinations of novel with established agents. Specifically, given the favorable characteristics of C1311, clinical evaluation of this agent alone or in combination with paclitaxel, for patients with metastatic bladder cancer that have failed first-line platinum therapy seems indicated.

Acknowledgments

The authors thank all members of the Theodorescu Laboratory for helpful thoughts and discussions, as well as Paul Williams, from the UVA Department of Public Health Sciences for his helpful suggestions.

References

[1] Botteman MF, Pashos CL, Redaelli A, Laskin B, and Hauser R (2003). The health economics of bladder cancer: a comprehensive review of the published literature. *Pharmacoeconomics* **21**, 1315–1330.

[2] Stein JP, Lieskovsky G, Cote R, Groshen S, Feng A-C, Boyd S, Skinner E, Bochner B, Thangathurai D, Mikhail M, et al. (2001). Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J Clin Oncol* **19**, 666–675.

[3] Gallagher DJ, Milowsky MI, and Bajorin DF (2008). Advanced bladder cancer: status of first-line chemotherapy and the search for active agents in the second-line setting. *Cancer* **113**, 1284–1293.

[4] Smith SC, Baras AS, Lee JK, and Theodorescu D (2010). The COXEN principle: translating signatures of *in vitro* chemosensitivity into tools for clinical outcome prediction and drug discovery in cancer. *Cancer Res* **70**, 1753–1758.

[5] Lee JK, Havaleshko DM, Cho H, Weinstein JN, Kaldjian EP, Karpovich J, Grimshaw A, and Theodorescu D (2007). A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. *Proc Natl Acad Sci USA* **104**, 13086–13091.

[6] Williams PD, Cheon S, Havaleshko DM, Jeong H, Cheng F, Theodorescu D, and Lee JK (2009). Concordant gene expression signatures predict clinical outcomes of cancer patients undergoing systemic therapy. *Cancer Res* **69**, 8302–8309.

[7] Shoemaker RH (2006). The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* **6**, 813–823.

[8] Berger B, Marquardt H, and Westendorf J (1996). Pharmacological and toxicological aspects of new imidazoacridinone antitumor agents. *Cancer Res* **56**, 2094–2104.

[9] Burger AM, Jenkins TC, Double JA, and Bibby MC (1999). Cellular uptake, cytotoxicity and DNA-binding studies of the novel imidazoacridinone antineoplastic agent C1311. *Br J Cancer* **81**, 367–375.

[10] Skladanowski A, Plisov SY, Konopa J, and Larsen AK (1996). Inhibition of DNA topoisomerase II by imidazoacridinones, new antineoplastic agents with strong activity against solid tumors. *Mol Pharmacol* **49**, 772–780.

[11] Chau M, Otake Y, Christensen JL, Fernandes DJ, and Ajami AM (2006). The imidazoacridinone, C-1311 (Symadex™): the first of a potent new class of FLT3 inhibitors. *AACR Meeting Abstracts* **2006**, B35.

[12] Otake Y, Chau M, Ajami A, and Fernandes D (2007). Mechanism of topoisomerase II inhibition by the imidazoacridinone, C-1311. *AACR Meeting Abstracts* **2007**, 4041.

[13] Thomas A, Anthoney A, Ahmed S, Drouin M, Major A, Capizzi RL, Grieshaber C, Loadman P, and Twelves C (2006). Evaluation of the safety of C-1311 administered in a phase 1 dose-escalation trial as a 1-hour infusion once every 3 weeks in patients with advanced solid tumors. *ASCO Meeting Abstracts* **24**, 12005.

[14] Thomas AL, Anthoney A, Scott E, Ahmed S, Lundberg AS, Major A, Capizzi RL, and Twelves CJ (2008). C-1311, a novel inhibitor of FLT3 and topoisomerase II: a phase 1 trial of a once every three week schedule in patients with advanced solid tumors. *J Clin Oncol* **26**, 2576.

[15] Havaleshko DM, Cho H, Conaway M, Owens CR, Hampton G, Lee JK, and Theodorescu D (2007). Prediction of drug combination chemosensitivity in human bladder cancer. *Mol Cancer Ther* **6**, 578–586.

[16] Hollingshead MG, Alley MC, Camalier RF, Abbott BJ, Mayo JG, Malspeis L, and Grever MR (1995). *In vivo* cultivation of tumor cells in hollow fibers. *Life Sci* **57**, 131–141.

[17] Pan X, Yuan DS, Ooi S-L, Wang X, Sookhai-Mahadeo S, Meluh P, and Boeke JD (2007). dSLAM analysis of genome-wide genetic interactions in *Saccharomyces cerevisiae*. *Methods* **41**, 206–221.

[18] Pierce SE, Fung EL, Jaramillo DF, Chu AM, Davis RW, Nislow C, and Giaever G (2006). A unique and universal molecular barcode array. *Nat Methods* **3**, 601–603.

[19] Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, and Sherlock G (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**, 3710–3715.

[20] Blower PE, Yang C, Fligner MA, Verducci JS, Yu L, Richman S, and Weinstein JN (2002). Pharmacogenomic analysis: correlating molecular substructure classes with microarray gene expression data. *Pharmacogenomics J* **2**, 259–271.

[21] Shankavaram UT, Reinhold WC, Nishizuka S, Major S, Morita D, Chary KK, Reimers MA, Scherf U, Kahn A, Dolginow D, et al. (2007). Transcript and protein expression profiles of the NCI-60 cancer cell panel: an integromic microarray study. *Mol Cancer Ther* **6**, 820–832.

[22] Weinstein JN, Myers TG, O'Connor PM, Friend SH, Fornace AJ Jr, Kohn KW, Fojo T, Bates SE, Rubinstein LV, Anderson NL, et al. (1997). An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**, 343–349.

[23] Shah MA and Schwartz GK (2001). Cell cycle–mediated drug resistance: an emerging concept in cancer therapy. *Clin Cancer Res* **7**, 2168–2181.

[24] Kaelin WG Jr (2005). The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* **5**, 689–698.

[25] Parsons AB, Brost RL, Ding H, Li Z, Zhang C, Sheikh B, Brown GW, Kane PM, Hughes TR, and Boone C (2004). Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat Biotechnol* **22**, 62–69.

[26] von der Maase H, Hansen SW, Roberts JT, Dogliotti L, Oliver T, Moore MJ, Bodrogi I, Albers P, Knuth A, Lippert CM, et al. (2000). Gemcitabine and cisplatin *versus* methotrexate, vinblastine, doxorubicin, and cisplatin in advanced or metastatic bladder cancer: results of a large, randomized, multinational, multicenter, phase III study. *J Clin Oncol* **18**, 3068–3077.

[27] Galsky MD (2005). The role of taxanes in the management of bladder cancer. *Oncologist* **10**, 792–798.

[28] Entwistle RA, Winefield RD, Foland TB, Lushington GH, and Himes RH (2008). The paclitaxel site in tubulin probed by site-directed mutagenesis of *Saccharomyces cerevisiae* β-tubulin. *FEBS Lett* **582**, 2467–2470.

[29] Simon R (2008). The use of genomics in clinical trial design. *Clin Cancer Res* **14**, 5984–5993.

[30] Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, and Cordon-Cardo C (2006). Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *J Clin Oncol* **24**, 778–789.

[31] Atiya AF (2005). Estimating the posterior probabilities using the *K*-nearest neighbor rule. *Neural Comput* **17**, 731–740.

[32] Als AB, Dyrskjot L, von der Maase H, Koed K, Mansilla F, Toldbod HE, Jensen JL, Ulhoi BP, Sengelov L, Jensen KM, et al. (2007). Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy in patients with advanced bladder cancer. *Clin Cancer Res* **13**, 4407–4414.

[33] Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, and Chinnaiyan AM (2004). ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* **6**, 1–6.

[34] Isambert N, Campone M, Bourbouloux E, Drouin M, Major A, Loadman P, Capizzi R, Grieshaber C, and Fumoleau P (2006). Evaluation of the safety of C-1311 administered in a phase I dose-escalation trial as a weekly infusion for 3 consecutive weeks in patients with advanced solid tumors. *ASCO Meeting Abstracts* **24**, 2069.

[35] Konopa JK, Koba M, and Dyrcz A (2005). Interstrand crosslinking of DNA by C-1311 (Symadex) and other imidazoacridinones. *AACR Meeting Abstracts* **2005**, 1382-a.

[36] Nguyen HB, Duncan K, Fang X, Chau M, Ajami A, and Small D (2008). FLT3 inhibition of leukemia cells by C-1311 and its analogs. *AACR Meeting Abstracts* **2008**, 4753.

[37] Capizzi RL, Roman LA, Tjulandin S, Smirnova I, Manikhas A, Paterson JS, Major A, Lundberg AS, and Fumoleau P (2008). Phase II trial of C1311, a novel inhibitor of topoisomerase II in advanced breast cancer. *J Clin Oncol* **26**, 1055.

Supplementary Methods

Hollow Fiber Assay

National Institutes of Health and University of Virginia ACUC guidelines were strictly observed. The National Cancer Institute HFA, developed by Hollingshead et al. [1] was performed as described to evaluate the *in vivo* activity of these two imidazoacridinone drugs and in combination with other chemotherapeutics in two physiologic compartments of the mouse, subcutaneous (SC) and intraperitoneal (IP). We used an *in vitro* control incubation to verify cell viability, sterility, and drug activity. Treated animals were compared to untreated controls with compounds administered daily on days 1 to 4 by IP injection. The compounds were administered once daily on day 1 to 4 by IP injection. Individual mouse body weights were recorded daily, and treatments would have been discontinued if an individual mouse body weight decreased ≥3 g or if other signs of toxicity/distress were evident, which did not occur. Imidazoacridinones were given at 20 mg/kg daily × 4, whereas paclitaxel was added at 15 mg/kg daily × 4. All mice were sacrificed on day 5, fibers were removed, and viable cell mass was quantified by the "stable end point" MTT dye conversion assay. Values presented are averages across four treated mice.

Analysis of Yeast TAG Array and Synthetic Lethal Data

The Affymetrix Yeast TAG4.0 array data were analyzed using software developed by the Giaever laboratory (http://chemogenomics.stanford.edu/supplements/04tag/), which quantile normalizes up and down tag intensities separately, applies quality filters, estimates and subtracts background from the treatment and control intensities, calculates $\log_2$ ratio of treatment over control enrichment, and identifies yeast strains that display significantly reduced fitness in a group of drug-treated replicates compared to control replicates. The method is detailed in the software documentation and companion publication [2]. The Yeast TAG4.0 drug data were converted to binary data as discussed in the Results section. Similarly, we converted the synthetic lethal data to binary data by assigning a 1 to synthetic lethal query-target pairs and 0 to all other ORF pairs. Combining the drug and synthetic lethal data resulted in a binary matrix with 1521 rows of yeast query genes and six drug treatments (four C1311 and two benomyl) and 2804 columns of yeast target genes. We note that the original size of each drug binary vector was 6431 (i.e., the number of yeast deletion strains interrogated on the array) and reduced to 2804 after being projected onto the set of available yeast target genes. We generated Figure 2, *A* to *C*, by applying two-way hierarchical clustering (i.e., cluster both rows and columns) with a cosine distance metric to this 1521 × 2804 binary matrix using the *clustergram* function in MATLAB Version 7.9.0 (The Mathworks, Natick, MA). Lists of yeast strains with reduced fitness for benomyl and C1311 were examined for statistically significant enrichment of gene ontology terms by GO::TermFinder [19] using default settings.

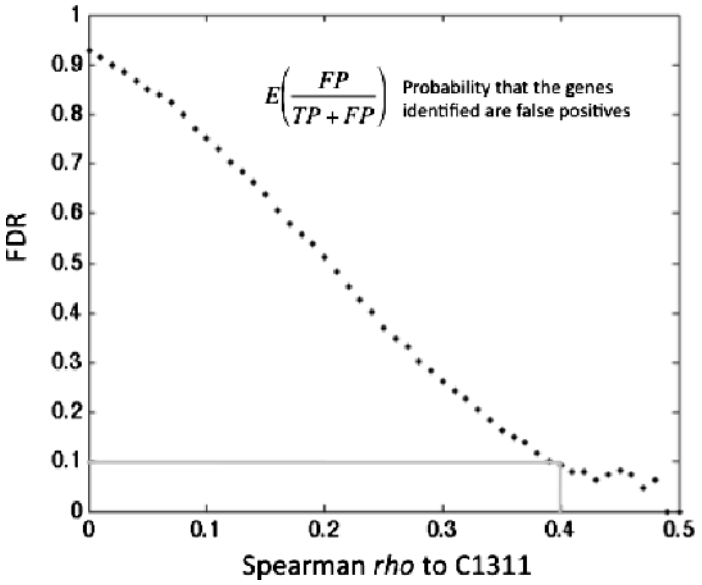
Development and Testing of a GEM Predictive of C1311 Sensitivity

Data sets Used:

1. All data sets used are Affymetrix HG-U133A and publicly available.
2. The BLA-40 data set is available as *GSE5845* at NCBI GEO (www.ncbi.nlm.nih.gov/geo).
3. The NCI-60 cell line data set is *GSE5720*, also on NCBI GEO.
4. The Sanchez-Carbayo et al. data set is available as supplementary data online with the referenced manuscript http://jco.ascopubs.org/cgi/content/full/24/5/778 [3]
5. The Als et al. data set [4] is *GSE5287*, also on NCBI GEO.
6. In all cases, authors' processed data were downloaded and used, log-transformed (if not already) and *z*-scored for standardization for inter–data set comparison.

Biomarker Discovery

The Spearman rank correlations of expression of each of the 22283 Affymetrix probes on the U133A platform across the NCI-60 cell line panels to the C1311 IC_{50} values for these cells were first calculated in Matlab (The Mathworks). To identify an appropriate cutoff point for these correlation values, we conducted random permutation testing to estimate the false discovery rate [5,6] at various cutoff values. We carried out 100 random permutation tests and recorded how many probes exhibited correlation values greater than the various cutoff points tested. Specifically, we examined absolute correlation values from 0.0 to 0.5 by 0.01 intervals, as shown:



By comparing the number of probes identified on average from the random permutation tests *versus* the number identified in the actual data across the range of absolute correlation values mentioned, we chose to accept a 10% false discovery rate rate, which represented a threshold correlation coefficient (ρ) of 0.40. These methods identified 219 probes that exhibited a significant correlation to C1311.

Development and Testing of the GEM

To help further refine these 219 probes and uncover subsets maintaining concordant expression between the two cell lines and human tumor data sets, we next used an application of the cross-correlation

step of the COXEN algorithm, however, adapted to multiple data sets. Three cross-correlation comparisons were made, namely 1) NCI-60 to BLA-40, 2) NCI-60 to S-C et al., and 3) BLA-40 to S-C et al. To select three-way concordant probes, we systematically examined a range of cross-correlation coefficient cutoff values, specifically 0.00 to 0.50 by 0.01 intervals. At each cutoff value, we recorded the set of genes exhibiting greater than threshold cross-correlation levels across *all three comparisons*. For *each* set of concordant probes, we then conducted the following procedures to assess their predictive performance in the BLA-40:

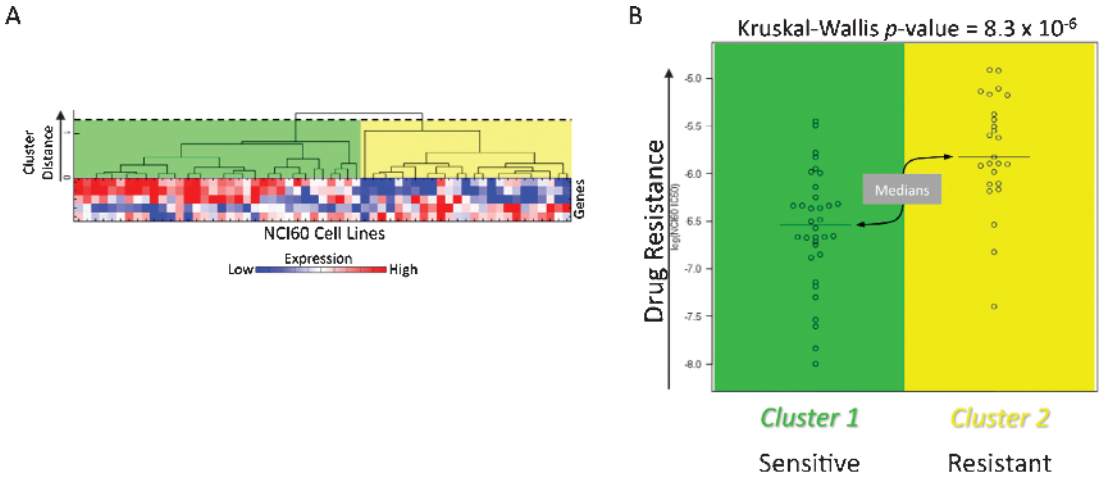
1) Selection of C1311-sensitive and -resistant NCI-60 training sets based on hierarchical clustering of cell lines across expression of C1311 sensitivity probes. Using a semisupervised approach, we discretized the continuous IC_{50} values in the NCI-60 data set into groupings of "sensitive" and "resistant" cells to provide categorical labels for the training data used in the development of a classifier. This was done by clustering the NCI-60 cell lines based on the expression of a given set of concordant C1311 sensitivity probes (described above) in an agglomerative hierarchical tree (e.g., see A in the following graph). Spearman correlation was used as the distance metric and the unweighted pair group method with arithmetic mean as the linkage function to construct the agglomerative hierarchical tree. Using the cluster function of Matlab, the hierarchical tree was used to divide the NCI-60 cells into two groups by drawing a horizontal cut through the tree such that only two clusters remain. This cluster grouping exhibited highly significant differences in NCI-60 IC_{50} values, as expected given our semisupervised approach (e.g., see B in the following graph). Examination of the central tendencies of the IC_{50} values from this grouping allowed us to appropriately label which cluster of the grouping represented sensitive (low IC_{50}) *versus* resistant (high IC_{50}) cell lines and use them for training data for the classifier.

were first log transformed and then *z*-score standardized to enable intercohort comparisons in correlation space. A weighted *k* nearest-neighbor (weighted kNN) algorithm [7,8] was used as the classifier, with the NCI-60 groupings from **1)** serving as the training data and predictions made for each cell line in the BLA-40 data set. A Spearman correlation distance metric was used to weight the influence that training samples had on the prediction of test samples, and the prediction of each test sample was based only on positively correlated training samples. The resulting predictions on the BLA-40 data set thus represent a binary classification of "sensitive" or "resistant," and we tested for difference in distributions of observed IC_{50} values for C1311 between predicted sensitive and resistant classes by nonparametric *t* tests.

Following these two procedures systematically for each of the 0.00 to 0.50 cross-correlation cutoffs as outlined above, we identified a five-gene model that best allowed us to differentiate resistant and sensitive BLA-40 cell lines based on the expression patterns in the NCI-60 data set (cross-correlation cutoff, or COXEN coefficient = 0.25). Importantly, the weighted kNN prediction algorithm used (Matlab Code available on request) also provides a posterior probability estimate for the classification call, a technique that has been reported before [9]. The program then uses a threshold of greater than or less than 0.5 as the threshold for binary classification as sensitive or resistant. We have termed this score the "C1311 sensitivity score," which ranges from 0 for sensitive to 1 for resistant, and it was the distributions of these scores for individual patient tumors that were tested among the different clinicopathologic characteristics in the data sets of Sanchez-Carbayo et al. and Als et al.

Significance of Best Classifier Performance by Random Permutation Testing

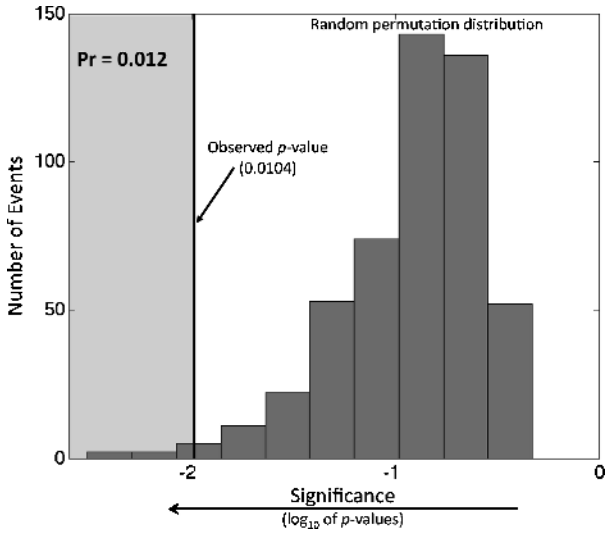
We next determined the exact statistical significance of these findings through permutation testing. To prove that the results we have



2) Evaluation of significance of predictions on the BLA-40. With these two groups of NCI-60 cells in hand, we next predicted which class (sensitive or resistant) each of the BLA-40 cell lines was most like. To do this, the NCI-60 and BLA-40 gene expression data

generated thus far cannot be ascribed to "overtraining" or random chance, we carefully carried out the identical procedures described before but with 219 randomly selected probes. We then repeated this random resampling test 500 times and recorded the *P* values for the

best-performing models, precisely as was done in **2)** with the 219 genes significantly associated to C1311 IC_{50} values in the NCI-60. This allowed us to estimate the *exact P* value that results similar to or better than those observed could be attributed to chance alone ($P = .012$), as shown in the following graph:



References

[1] Hollingshead MG, Alley MC, Camalier RF, Abbott BJ, Mayo JG, Malspeis L, and Grever MR (1995). *In vivo* cultivation of tumor cells in hollow fibers. *Life Sci* **57**, 131–141.

[2] Pierce SE, Fung EL, Jaramillo DF, Chu AM, Davis RW, Nislow C, and Giaever G (2006). A unique and universal molecular barcode array. *Nat Methods* **3**, 601–603.

[3] Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, and Cordon-Cardo C (2006). Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *J Clin Oncol* **24**, 778–789.

[4] Als AB, Dyrskjot L, von der Maase H, Koed K, Mansilla F, Toldbod HE, Jensen JL, Ulhoi BP, Sengelov L, Jensen KM, et al. (2007). Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy in patients with advanced bladder cancer. *Clin Cancer Res* **13**, 4407–4414.

[5] Farcomeni A (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Stat Methods Med Res* **17**, 347–388.

[6] Pawitan Y, Michiels S, Koscielny S, Gusnanto A, and Ploner A (2005). False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21**, 3017–3024.

[7] Berrar D, Bradbury I, and Dubitzky W (2006). Instance-based concept learning from multiclass DNA microarray data. *BMC Bioinformatics* **7**, 73.

[8] Horlings HM, van Laar RK, Kerst J-M, Helgason HH, Wesseling J, van der Hoeven JJM, Warmoes MO, Floore A, Witteveen A, Lahti-Domenici J, et al. (2008). Gene expression profiling to identify the histogenetic origin of metastatic adenocarcinomas of unknown primary. *J Clin Oncol* **26**, 4435–4441.

[9] Atiya AF (2005). Estimating the posterior probabilities using the K-nearest neighbor rule. *Neural Comput* **17**, 731–740.

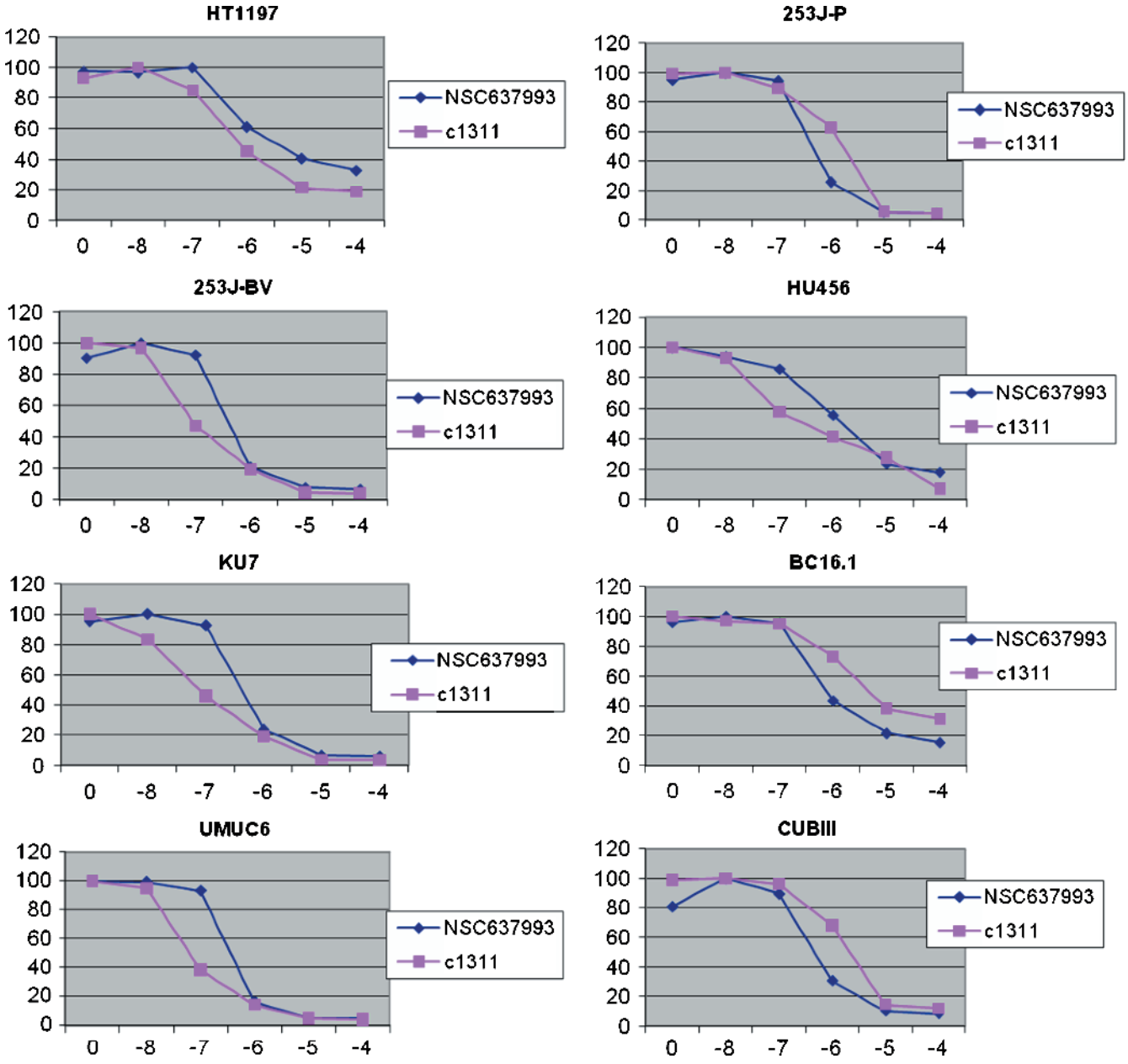


Figure W1. C1311 and NSC-637993 dose-response curves. Cells were exposed to RPMI 1640/10% FBS medium with diluted NSC-637993 or C1311 at concentrations of 0 (control), 0.01, 0.1, 1, 10, and 100 μM; after 72 hours of culture, cell counts were assayed. Each concentration of drug was tested on six replicate wells in more than 4 individual experiments. Data presented below for each of the cell lines show percent of maximal cell growth (*y* axis) per drug, averaged across the four replicates, plotted against the $\log_{10}$ treatment dose. These drug-response curves were used in Spline regression to estimate the IC_{50} values presented in Figure 1*A* and Table W1 as described in Materials and Methods.

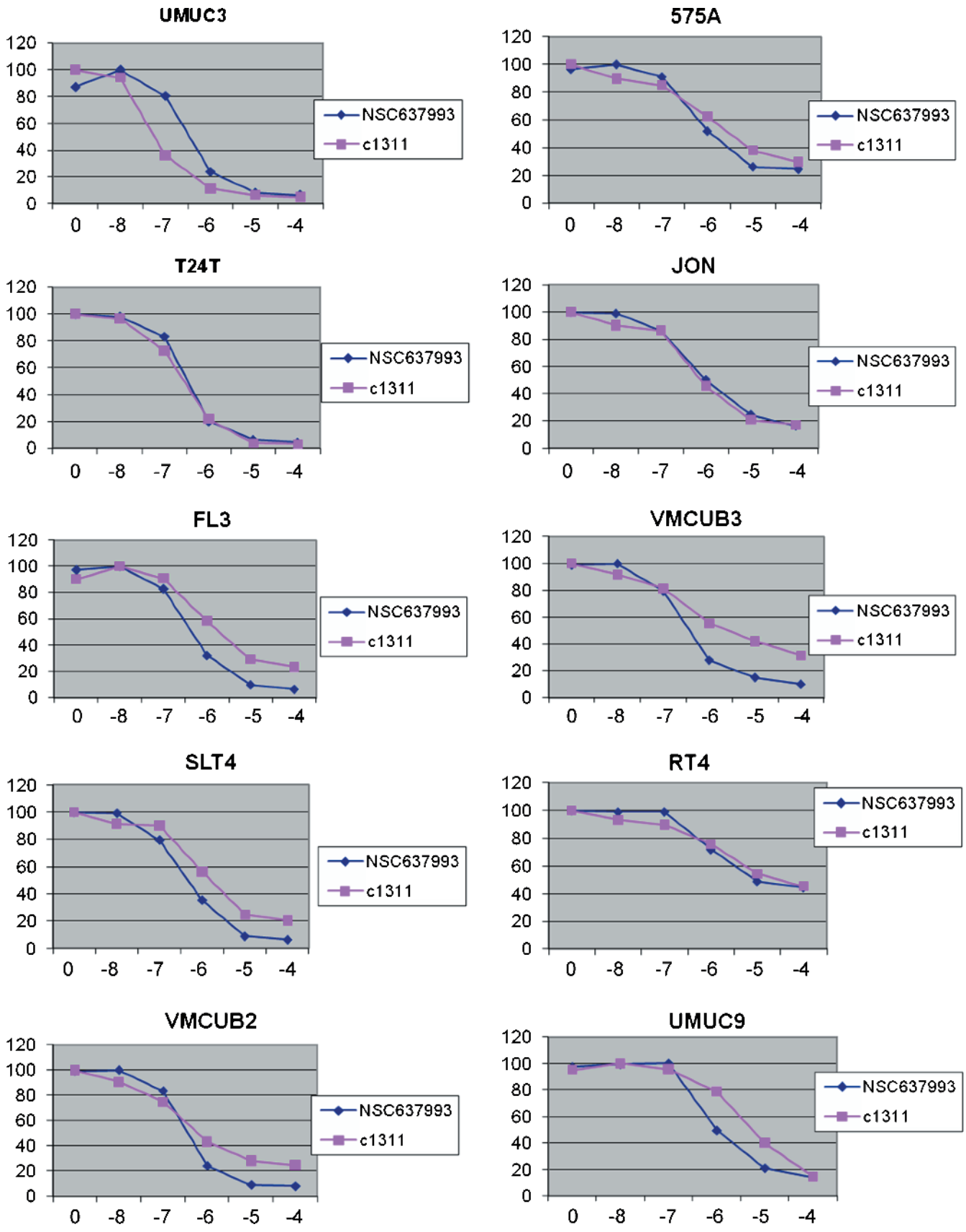


Figure W1. (*continued*).

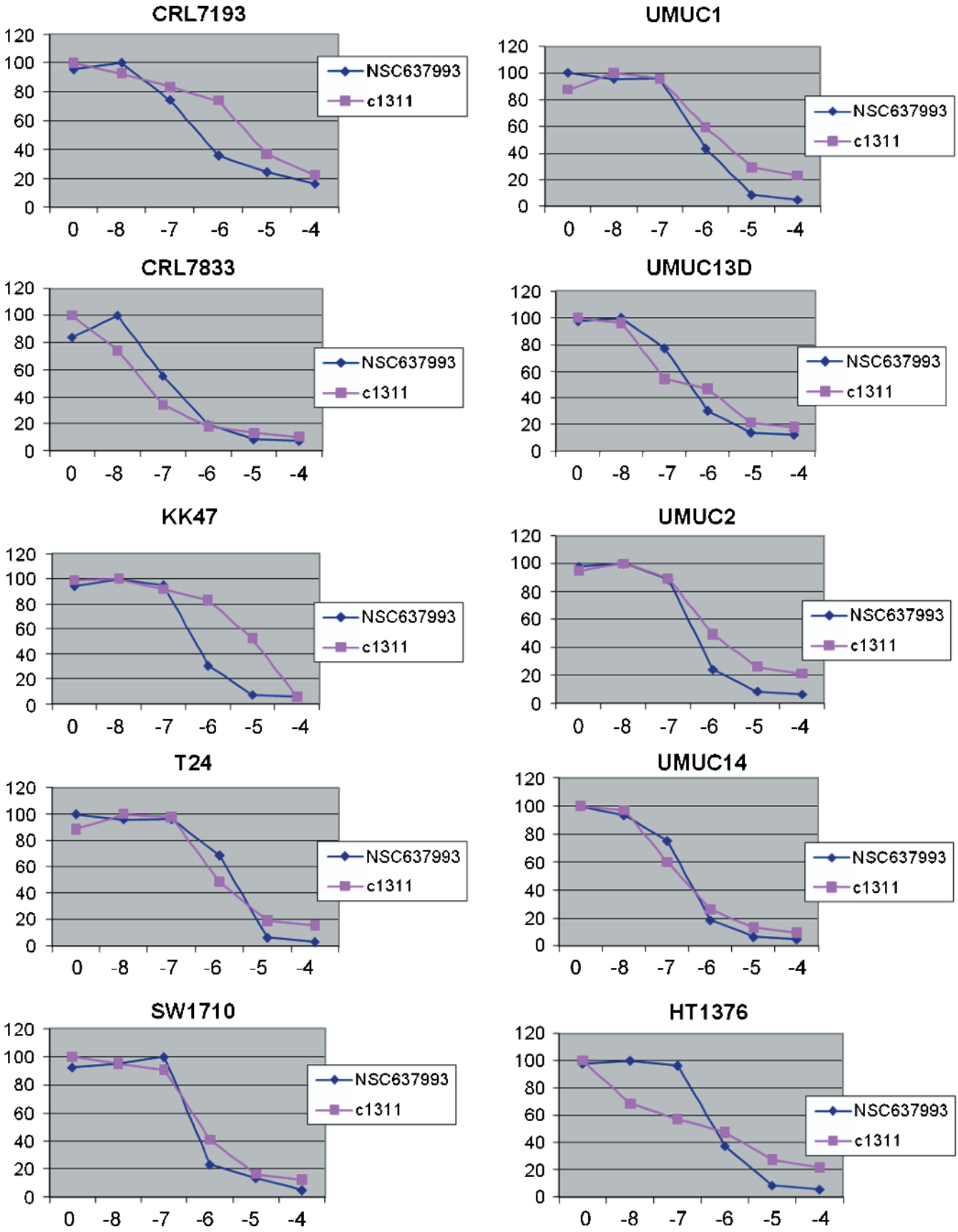


Figure W1. (*continued*).

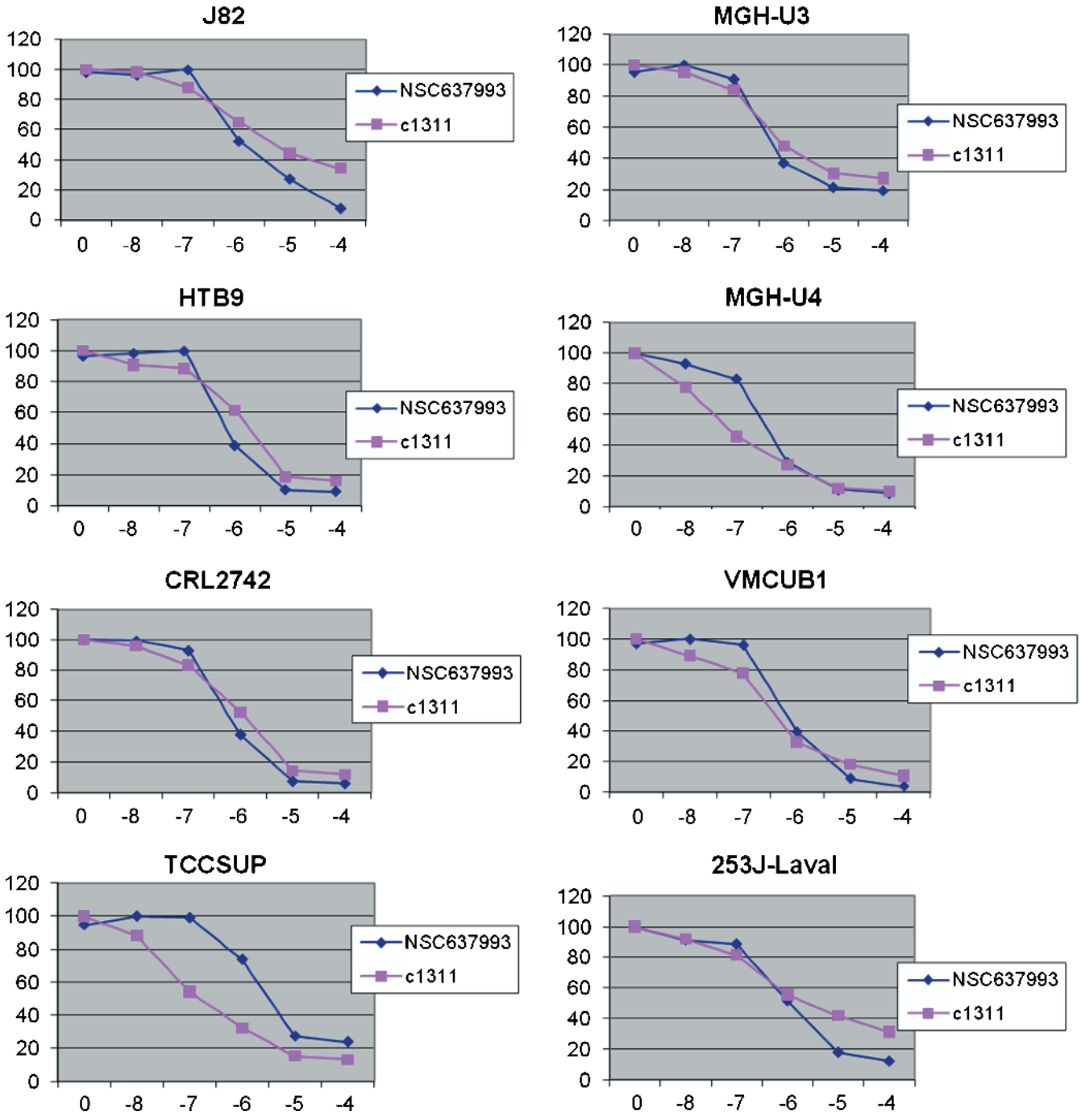


Figure W1. (*continued*).

Table W1. Log_{10} IC_{50} Values for the BLA-40 Cell Line Panel.

Cell Line*	C1311	NSC-637993†
CRL7833	−7.173	−6.574
KU7	−6.954	−6.378
UMUC3	−6.951	−6.337
UMUC6	−6.923	−6.281
MGH-U4	−6.874	−6.187
253J-BV	−6.836	−6.342
TCCSUP	−6.682	−5.193
SCaBER	−6.671	‡
UMUC14	−6.651	−6.641
T24T	−6.622	−6.604
HU456	−6.476	−5.675
UMUC3-E	−6.454	‡
VMCUB1	−6.383	−6.071
UMUC13D	−6.353	−6.361
VMCUB2	−6.197	−6.433
HT1376	−6.137	−5.959
CRL2169	−6.107	‡
SW1710	−6.101	−6.238
JON	−5.972	−5.926
MGH-U3	−5.965	−6.033
HT1197	−5.949	−5.245
CRL2742	−5.893	−6.135
UMUC2	−5.781	−6.425
253J-P	−5.727	−6.320
SLT4	−5.676	−6.273
T24	−5.656	−5.610
HTB9	−5.645	−5.946
253JLaval	−5.597	−5.722
VMCUB3	−5.597	−6.269
CUBIII	−5.565	−6.075
FL3	−5.448	−6.404
575A	−5.418	−5.634
UMUC1	−5.351	−6.009
CRL7193	−5.327	−6.327
J82	−5.269	−5.648
PSI	−5.266	‡
BC16.1	−5.231	−5.930
UMUC9	−5.180	−5.751
KK47	−4.841	−6.203
RT4	−4.722	−4.672

*The BLA-40 cell line panel has been reported before [15].
†Reported IC_{50} values for NSC-637993 [15] were recalculated using Spline regression and listed here for comparison between the related imidazoacridinones.
‡Four of the BLA-40 cell lines were not tested for NSC-637993 in the prior report.

Table W2. HFA Results for NSC-637993 (A) and NSC and Paclitaxel (B).

(A)

Tx	Cell Line	NSC Log_{10} IC_{50}	SC*	P†	IP‡	P†	Overall§	Overall P†
NSC Alone	T24T	−6.60	62.7	.005	71.6	.0006	61.3	<.0001
	FL3	−6.40	66.0	.0002	56.6	<.0001	67.2	<.0001
	UMUC1	−6.00	88.9	.2	54.9	<.0001	71.9	.0115
	KK47	−6.20	118.0	.04	107.8	.0002	112.8	.0018

(B)

Tx	Cell Line	NSC Log_{10} IC_{50}	Paclitaxel Log_{10} IC_{50}¶	SQ*	P†	IP‡	P†	Overall	Overall P†
NSC + Paclitaxel	UMUC6	−6.28	−8.99	12.5	<.0001	27.0	<.0001	19.8	<.0001
	HTB9	−5.94	−8.93	20.6	<.0001	8.1	<.0001	14.4	.0006
	T24	−5.61	−7.36	62.8	<.0001	28.3	<.0001	45.6	<.0001
	KK47	−6.20	>−7	94.3	.04	37.1	<.0001	65.7	.02

*Average percentage of control growth across four replicates at the subcutaneous implantation site.
†Two-tailed P value for single-sample t test against the hypothesis that the inhibition was 0%.
‡Average percentage of control growth across four replicates at the intraperitoneal implantation site.
§Average percentage of control growth across all replicates and sites.
¶We have reported *in vitro* sensitivities to paclitaxel across the BLA-40 panel before [16].

Table W3. Union of Yeast Strains with Reduced Fitness in C1311 (A) and Benomyl (B) Treatment.

Symbol	Description
(A)	
ACB1	Acyl-CoA-binding protein, transports newly synthesized acyl-CoA esters from fatty acid synthetase (Fas1p-Fas2p) to acyl-CoA-consuming processes
ACE2	Transcription factor that activates expression of early G_1-specific genes, localizes to daughter cell nuclei after cytokinesis and delays G_1 progression in daughters, localization is regulated by phosphorylation; potential Cdc28p substrate
ADO1	Adenosine kinase, required for the utilization of *S*-adenosylmethionine (AdoMet); may be involved in recycling adenosine produced through the methyl cycle
AIM22	Putative lipoate-protein ligase, required along with Lip2 and Lip5 for lipoylation of Lat1p and Kgd2p; similar to *E. coli* LplA; null mutant displays reduced frequency of mitochondrial genome loss
AKR1	Palmitoyl transferase involved in protein palmitoylation; acts as a negative regulator of pheromone response pathway; required for endocytosis of pheromone receptors; involved in cell shape control; contains ankyrin repeats
ALG13	Catalytic component of UDP-GlcNAc transferase, required for the second step of dolichyl-linked oligosaccharide synthesis; anchored to the ER membrane through interaction with Alg14p; similar to bacterial and human glycosyltransferases
ARC15	Subunit of the ARP2/3 complex, which is required for the motility and integrity of cortical actin patches
ARO1	Pentafunctional arom protein, catalyzes steps 2 through 6 in the biosynthesis of chorismate, which is a precursor to aromatic amino acids
ARO2	Bifunctional chorismate synthase and flavin reductase, catalyzes the conversion of 5-enolpyruvylshikimate 3-phosphate (EPSP) to form chorismate, which is a precursor to aromatic amino acids
BEM2	Rho GTPase activating protein (RhoGAP) involved in the control of cytoskeleton organization and cellular morphogenesis; required for bud emergence
BEM4	Protein involved in establishment of cell polarity and bud emergence; interacts with the Rho1p small GTP–binding protein and with the Rho-type GTPase Cdc42p; involved in maintenance of proper telomere length
BIM1	Microtubule-binding protein that together with Kar9p makes up the cortical microtubule capture site and delays the exit from mitosis when the spindle is oriented abnormally
BUD20	Protein involved in bud site selection; diploid mutants display a random budding pattern instead of the wild-type bipolar pattern
COY1	Golgi membrane protein with similarity to mammalian CASP; genetic interactions with GOS1 (encoding a Golgi snare protein) suggest a role in Golgi function
CPA1	Small subunit of carbamoyl phosphate synthetase, which catalyzes a step in the synthesis of citrulline, an arginine precursor; translationally regulated by an attenuator peptide encoded by YOR302W within the CPA1 mRNA 5′-leader
CSR1	Phosphatidylinositol transfer protein with a potential role in regulating lipid and fatty acid metabolism under heme-depleted conditions; interacts specifically with thioredoxin peroxidase; may have a role in oxidative stress resistance
CTF4	Chromatin-associated protein, required for sister chromatid cohesion; interacts with DNA polymerase α (Pol1p) and may link DNA synthesis to sister chromatid cohesion
CYB5	Cytochrome b5, involved in the sterol and lipid biosynthesis pathways; acts as an electron donor to support sterol C5-6 desaturation
DYN1	Cytoplasmic heavy chain dynein, microtubule motor protein, required for anaphase spindle elongation; involved in spindle assembly, chromosome movement, and spindle orientation during cell division, targeted to microtubule tips by Pac1p
ERG2	C-8 sterol isomerase, catalyzes the isomerization of the delta-8 double bond to the delta-7 position at an intermediate step in ergosterol biosynthesis
ERG25	C-4 methyl sterol oxidase, catalyzes the first of three steps required to remove two C-4 methyl groups from an intermediate in ergosterol biosynthesis; mutants accumulate the sterol intermediate 4,4-dimethylzymosterol
ERG28	Endoplasmic reticulum membrane protein, may facilitate protein-protein interactions between the Erg26p dehydrogenase and the Erg27p 3-ketoreductase and/or tether these enzymes to the ER, also interacts with Erg6p
ERG3	C-5 sterol desaturase, catalyzes the introduction of a C-5(6) double bond into episterol, a precursor in ergosterol biosynthesis; mutants are viable, but cannot grow on nonfermentable carbon sources
ERG6	Delta(24)-sterol C-methyltransferase, converts zymosterol to fecosterol in the ergosterol biosynthetic pathway by methylating position C-24; localized to both lipid particles and mitochondrial outer membrane
FEN2	Plasma membrane H^+-pantothenate symporter; confers sensitivity to the antifungal agent fenpropimorph
GCR2	Transcriptional activator of genes involved in glycolysis; interacts and functions with the DNA-binding protein Gcr1p
GET2	Subunit of the GET complex; involved in insertion of proteins into the ER membrane; required for the retrieval of HDEL proteins from the Golgi to the ER in an ERD2 dependent fashion and for meiotic nuclear division
GLO4	Mitochondrial glyoxalase II, catalyzes the hydrolysis of *S*-D-lactoylglutathione into glutathione and D-lactate
GTR2	Putative GTP binding protein that negatively regulates Ran/Tc4 GTPase cycle; activates transcription; subunit of EGO and GSE complexes; required for sorting of Gap1p; localizes to cytoplasm and to chromatin; homolog of human RagC and RagD
ISC1	Mitochondrial membrane localized inositol phosphosphingolipid phospholipase C, hydrolyzes complex sphingolipids to produce ceramide; activated by phosphatidylserine, cardiolipin, and phosphatidylglycerol; mediates Na^+ and Li^+ halotolerance
MGA2	ER membrane protein involved in regulation of OLE1 transcription, acts with homolog Spt23p; inactive ER form dimerizes and one subunit is then activated by ubiquitin/proteasome–dependent processing followed by nuclear targeting
MGM1	Mitochondrial GTPase related to dynamin, present in a complex containing Ugo1p and Fzo1p; required for normal morphology of cristae and for stability of Tim11p; homolog of human OPA1 involved in autosomal dominant optic atrophy
MNN11	Subunit of a Golgi mannosyltransferase complex that also contains Anp1p, Mnn9p, Mnn10p, and Hoc1p, and mediates elongation of the polysaccharide mannan backbone; has homology to Mnn10p
MNN9	Subunit of Golgi mannosyltransferase complex also containing Anp1p, Mnn10p, Mnn11p, and Hoc1p that mediates elongation of the polysaccharide mannan backbone; forms a separate complex with Van1p that is also involved in backbone elongation
MRPL16	Mitochondrial ribosomal protein of the large subunit
MSF1	Mitochondrial phenylalanyl-tRNA synthetase, active as a monomer, unlike the cytoplasmic subunit which is active as a dimer complexed to a β-subunit dimer; similar to the α subunit of *E. coli* phenylalanyl-tRNA synthetase
NPT1	Nicotinate phosphoribosyltransferase, acts in the salvage pathway of NAD^+ biosynthesis; required for silencing at rDNA and telomeres and has a role in silencing at mating-type loci; localized to the nucleus
PDX3	Pyridoxine (pyridoxamine) phosphate oxidase, has homologs in *E. coli* and *Myxococcus xanthus*; transcription is under the general control of nitrogen metabolism
PER1	Protein of the endoplasmic reticulum, required for GPI-phospholipase A2 activity that remodels the GPI anchor as a prerequisite for association of GPI-anchored proteins with lipid rafts; functionally complemented by human ortholog PERLD1
PFY1	Profilin, actin- and phosphatidylinositol 4,5-bisphosphate-binding protein, involved in cytoskeleton organization, required for normal timing of actin polymerization in response to thermal stress; localizes to plasma membrane and cytosol
PGA1	Essential component of GPI-mannosyltransferase II, responsible for second mannose addition to GPI precursors as a partner of Gpi18p; required for maturation of Gas1p and Pho8p; has synthetic genetic interactions with secretory pathway genes
PIB2	Protein binding phosphatidylinositol 3-phosphate, involved in telomere-proximal repression of gene expression; similar to Fab1 and Vps27
PPM1	Carboxyl methyltransferase, methylates the C-terminus of the protein phosphatase 2A catalytic subunit (Pph21p or Pph22p), which is important for complex formation with regulatory subunits
PRO1	γ-Glutamyl kinase, catalyzes the first step in proline biosynthesis
RAD51	Strand exchange protein, forms a helical filament with DNA that searches for homology; involved in the recombinational repair of double-strand breaks in DNA during vegetative growth and meiosis; homolog of Dmc1p and bacterial RecA protein

Table W3. (*continued*)

Symbol	Description
RAD55	Protein that stimulates strand exchange by stabilizing the binding of Rad51p to single-stranded DNA; involved in the recombinational repair of double-strand breaks in DNA during vegetative growth and meiosis; forms heterodimer with Rad57p
RAX2	N-glycosylated protein involved in the maintenance of bud site selection during bipolar budding; localization requires Rax1p; RAX2 mRNA stability is regulated by Mpt5p
RET2	Delta subunit of the coatomer complex (COPI), which coats Golgi-derived transport vesicles; involved in retrograde transport between Golgi and ER
RIM2	Mitochondrial pyrimidine nucleotide transporter; imports pyrimidine nucleoside triphosphates and exports pyrimidine nucleoside monophosphates; member of the mitochondrial carrier family
RML2	Mitochondrial ribosomal protein of the large subunit, has similarity to *E. coli* L2 ribosomal protein; fat21 mutant allele causes inability to use oleate and may interfere with activity of the Adr1p transcription factor
RPA135	RNA polymerase I subunit A135
RPB4	RNA polymerase II subunit B32; forms two subunit dissociable complex with Rpb7p; involved recruitment of 3′-end processing factors to transcribing RNA polymerase II complex and in export of mRNA to cytoplasm under stress conditions
RPD3	Histone deacetylase; regulates transcription and silencing; plays a role in regulating Ty1 transposition
RPL13B	Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl13Ap; not essential for viability; has similarity to rat L13 ribosomal protein
RPL1A	N-terminally acetylated protein component of the large (60S) ribosomal subunit, nearly identical to Rpl1Bp and has similarity to *E. coli* L1 and rat L10a ribosomal proteins; rpl1a rpl1b double null mutation is lethal /// N-terminally acetylated protein component of the large (60S) ribosomal subunit, nearly identical to Rpl1Ap and has similarity to *E. coli* L1 and rat L10a ribosomal proteins; rpl1a rpl1b double null mutation is lethal
RPL2B	Protein component of the large (60S) ribosomal subunit, identical to Rpl2Ap and has similarity to *E. coli* L2 and rat L8 ribosomal proteins; expression is upregulated at low temperatures
RPL35A	Protein component of the large (60S) ribosomal subunit, identical to Rpl35Bp and has similarity to rat L35 ribosomal protein /// Protein component of the large (60S) ribosomal subunit, identical to Rpl35Ap and has similarity to rat L35 ribosomal protein
RPL39	Protein component of the large (60S) ribosomal subunit, has similarity to rat L39 ribosomal protein; required for ribosome biogenesis; loss of both Rpl31p and Rpl39p confers lethality; also exhibits genetic interactions with SIS1 and PAB1
RPS14A	Ribosomal protein 59 of the small subunit, required for ribosome assembly and 20S pre-rRNA processing; mutations confer cryptopleurine resistance; nearly identical to Rps14Bp and similar to *E. coli* S11 and rat S14 ribosomal proteins
RPS1B	Ribosomal protein 10 (rp10) of the small (40S) subunit; nearly identical to Rps1Ap and has similarity to rat S3a ribosomal protein
RPS25A	Protein component of the small (40S) ribosomal subunit; nearly identical to Rps25Bp and has similarity to rat S25 ribosomal protein
RSB1	Suppressor of sphingoid long chain base (LCB) sensitivity of an LCB-lyase mutation; putative integral membrane transporter or flippase that may transport LCBs from the cytoplasmic side toward the extracytoplasmic side of the membrane
SAC1	Phosphatidylinositol phosphate (PtdInsP) phosphatase involved in hydrolysis of PtdIns[4]P; transmembrane protein localizes to ER and Golgi; involved in protein trafficking and processing, secretion, and cell wall maintenance
SEC59	Dolichol kinase, catalyzes the terminal step in dolichyl monophosphate (Dol-P) biosynthesis; required for viability and for normal rates of lipid intermediate synthesis and protein N-glycosylation
SGO1	Component of the spindle checkpoint, involved in sensing lack of tension on mitotic chromosomes; protects centromeric Rec8p at meiosis I; required for accurate chromosomal segregation at meiosis II and for mitotic chromosome stability
SGS1	Nucleolar DNA helicase of the RecQ family involved in genome integrity maintenance; regulates chromosome synapsis and meiotic joint molecule/crossover formation; similar to human BLM and WRN proteins implicated in Bloom and Werner syndromes
SMC3	Subunit of the multiprotein cohesin complex required for sister chromatid cohesion in mitotic cells; also required, with Rec8p, for cohesion and recombination during meiosis; phylogenetically conserved SMC chromosomal ATPase family member
SNF1	AMP-activated serine/threonine protein kinase found in a complex containing Snf4p and members of the Sip1p/Sip2p/Gal83p family; required for transcription of glucose-repressed genes, thermotolerance, sporulation, and peroxisome biogenesis
SOK2	Nuclear protein that plays a regulatory role in the cyclic AMP (cAMP)–dependent protein kinase (PKA) signal transduction pathway; negatively regulates pseudohyphal differentiation; homologous to several transcription factors
SOV1	Mitochondrial protein of unknown function
SPO7	Putative regulatory subunit of Nem1p-Spo7p phosphatase holoenzyme, regulates nuclear growth by controlling phospholipid biosynthesis, required for normal nuclear envelope morphology, premeiotic replication, and sporulation
SPT14	UDP-GlcNAc-binding and catalytic subunit of the enzyme that mediates the first step in glycosylphosphatidylinositol (GPI) biosynthesis, mutations cause defects in transcription and in biogenesis of cell wall proteins
SSQ1	Mitochondrial hsp70-type molecular chaperone, required for assembly of iron/sulfur clusters into proteins at a step after cluster synthesis, and for maturation of Yfh1p, which is a homolog of human frataxin implicated in Friedreich's ataxia
SUR4	Elongase, involved in fatty acid and sphingolipid biosynthesis; synthesizes very long chain 20-26-carbon fatty acids from C18-CoA primers; involved in regulation of sphingolipid biosynthesis
TBF1	Telobox-containing general regulatory factor; binds to TTAGGG repeats within subtelomeric anti–silencing regions (STARs) and possibly throughout the genome and mediates their insulating capacity by blocking silent chromatin propagation
TIP20	Peripheral membrane protein required for fusion of COPI vesicles with the ER, prohibits back fusion of COPII vesicles with the ER, may act as a sensor for vesicles at the ER membrane; interacts with Sec20p
TMA23	Nucleolar protein of unknown function implicated in ribosome biogenesis; TMA23 may be a fungal-specific gene as no homologs have been yet identified in higher eukaryotes
TRP3	Bifunctional enzyme exhibiting both indole-3-glycerol-phosphate synthase and anthranilate synthase activities, forms multifunctional hetero-oligomeric anthranilate synthase:indole-3-glycerol phosphate synthase enzyme complex with Trp2p
VAC7	Integral vacuolar membrane protein involved in vacuole inheritance and morphology; activates Fab1p kinase activity under basal conditions and also after hyperosmotic shock
VPS65	Protein coding
VPS8	Membrane-associated protein that interacts with Vps21p to facilitate soluble vacuolar protein localization; component of the CORVET complex; required for localization and trafficking of the CPY sorting receptor; contains RING finger motif
YIP1	Integral membrane protein required for the biogenesis of ER-derived COPII transport vesicles; interacts with Yif1p and Yos1p; localizes to the Golgi, the ER, and COPII vesicles
YML012C-A	Hypothetical ORF
YMR290W-A	Protein required for cell viability
YOR199W	Hypothetical ORF
YPL080C	Hypothetical ORF
YIL014C-A	Putative protein of unknown function
YBR056W	Putative cytoplasmic protein of unknown function
YBR221W-A	Putative protein of unknown function; identified by expression profiling and mass spectrometry
YDL118W	Nonessential protein of unconfirmed function; mutants are defective in telomere maintenance, and are synthetically sick or lethal with α-synuclein
YDL119C	Putative mitochondrial transport protein; GFP fusion protein is induced in response to the DNA-damaging agent MMS; the authentic, nontagged protein is detected in purified mitochondria

Table W3. (*continued*)

Symbol	Description
(B)	
ABD1	Methyltransferase, catalyzes the transfer of a methyl group from *S*-adenosylmethionine to the GpppN terminus of capped mRNA
AKR1	Palmitoyl transferase involved in protein palmitoylation; acts as a negative regulator of pheromone response pathway; required for endocytosis of pheromone receptors; involved in cell shape control; contains ankyrin repeats
ARP6	Actin-related protein that binds nucleosomes; a component of the SWR1 complex, which exchanges histone variant H2AZ (Htz1p) for chromatin-bound histone H2A
BLM10	Proteasome activator subunit; found in association with core particles, with and without the 19S regulatory particle; required for resistance to bleomycin, may be involved in protecting against oxidative damage; similar to mammalian PA200
BUB3	Kinetochore checkpoint WD40 repeat protein that localizes to kinetochores during prophase and metaphase, delays anaphase in the presence of unattached kinetochores; forms complexes with Mad1p-Bub1p and with Cdc20p, binds Mad2p and Mad3p
CIN1	Tubulin folding factor D involved in β-tubulin (Tub2p) folding; isolated as mutant with increased chromosome loss and sensitivity to benomyl
CIN2	GTPase-activating protein (GAP) for Cin4p; tubulin folding factor C involved in β-tubulin (Tub2p) folding; mutants display increased chromosome loss and benomyl sensitivity; deletion complemented by human GAP, retinitis pigmentosa 2
CIN4	GTP-binding protein involved in β-tubulin (Tub2p) folding; isolated as mutant with increased chromosome loss and sensitivity to benomyl; regulated by the GTPase-activating protein, Cin2p, the human retinitis pigmentosa 2 (RP2) homolog
CMC1	Evolutionarily conserved copper-binding protein of the mitochondrial intermembrane space, may be involved in delivering copper from the matrix to the cytochrome *c* oxidase complex; contains a twin CX9C motif
DAL82	Positive regulator of allophanate-inducible genes; binds a dodecanucleotide sequence upstream of all genes that are induced by allophanate; contains an UISALL DNA-binding, a transcriptional activation, and a coiled-coil domain
ECM23	Nonessential protein of unconfirmed function; affects pre-rRNA processing, may act as a negative regulator of the transcription of genes involved in pseudohyphal growth; homologous to Srd1p
ERG2	C-8 sterol isomerase, catalyzes the isomerization of the delta-8 double bond to the delta-7 position at an intermediate step in ergosterol biosynthesis
GIM3	Subunit of the heterohexameric cochaperone prefoldin complex which binds specifically to cytosolic chaperonin and transfers target proteins to it
GIM4	Subunit of the heterohexameric cochaperone prefoldin complex which binds specifically to cytosolic chaperonin and transfers target proteins to it
GND1	6-Phosphogluconate dehydrogenase (decarboxylating), catalyzes an NADPH regenerating reaction in the pentose phosphate pathway; required for growth on D-glucono-delta-lactone and adaptation to oxidative stress
MGR3	Subunit of the mitochondrial (mt) i-AAA protease supercomplex, which degrades misfolded mitochondrial proteins; forms a subcomplex with Mgr1p that binds to substrates to facilitate proteolysis; required for growth of cells lacking mtDNA
MTQ2	*S*-adenosylmethionine-dependent methyltransferase of the seven β-strand family; subunit of complex with Trm112p that methylates translation release factor Sup45p (eRF1) in the ternary complex eRF1-eRF3-GTP; similar to *E. coli* PrmC
PAC10	Part of the heteromeric co-chaperone GimC/prefoldin complex, which promotes efficient protein folding
PAC2	Microtubule effector required for tubulin heterodimer formation, binds α-tubulin, required for normal microtubule function, null mutant exhibits cold-sensitive microtubules and sensitivity to benomyl
PFD1	Subunit of heterohexameric prefoldin, which binds cytosolic chaperonin and transfers target proteins to it; involved in the biogenesis of actin and of α- and γ-tubulin
RVS161	Amphiphysin-like lipid raft protein; interacts with Rvs167p and regulates polarization of the actin cytoskeleton, endocytosis, cell polarity, cell fusion and viability following starvation or osmotic stress
SET6	SET domain protein of unknown function; deletion heterozygote is sensitive to compounds that target ergosterol biosynthesis, may be involved in compound availability
STB5	Transcription factor, involved in regulating multidrug resistance and oxidative stress response; forms a heterodimer with Pdr1p; contains a Zn(II)2Cys6 zinc finger domain that interacts with a pleiotropic drug resistance element *in vitro*
SWR1	Swi2/Snf2–related ATPase that is the structural component of the SWR1 complex, which exchanges histone variant H2AZ (Htz1p) for chromatin-bound histone H2A
TIF1	translation initiation factor eIF4E, 4F complex subunit (PMID 8955119)
TUB3	Alpha-tubulin; associates with β-tubulin (Tub2p) to form tubulin dimer, which polymerizes to form microtubules; expressed at lower level than Tub1p
UGO1	Protein of unknown function; outer membrane component of the mitochondrial fusion machinery; Ugo1p bind directly to Fzo1p and Mgm1p and thereby link these two GTPases during mitochondrial fusion
VPS53	Component of the GARP (Golgi-associated retrograde protein) complex, Vps51p-Vps52p-Vps53p-Vps54p, which is required for the recycling of proteins from endosomes to the late Golgi; required for vacuolar protein sorting
VPS71	Nucleosome-binding component of the SWR1 complex, which exchanges histone variant H2AZ (Htz1p) for chromatin-bound histone H2A; required for vacuolar protein sorting
VPS72	Htz1p-binding component of the SWR1 complex, which exchanges histone variant H2AZ (Htz1p) for chromatin-bound histone H2A; required for vacuolar protein sorting
YAF9	Subunit of both the NuA4 histone H4 acetyltransferase complex and the SWR1 complex, may function to antagonize silencing near telomeres; interacts directly with Swc4p, has homology to human leukemogenic protein AF9, contains a YEATS domain
YKE2	Subunit of the heterohexameric Gim/prefoldin protein complex involved in the folding of α-tubulin, β-tubulin, and actin
YLR269C	Hypothetical ORF
YML094C-A	Hypothetical ORF

Table W4. C1311 IC_{50}-Correlated Microarray Probes.

Probe Set*	Rho†	Symbol	Gene Title	Entrez Gene ID	Cytoband
200661_at	0.4553	*CTSA*	cathepsin A	5476	20q13.1
200677_at	0.4239	*PTTG1IP*	pituitary tumor-transforming 1 interacting protein	754	21q22.3
200696_s_at	0.5067	*GSN*	gelsolin (amyloidosis, Finnish type)	2934	9q33
200827_at	0.4218	*PLOD1*	procollagen-lysine 1, 2-oxoglutarate 5-dioxygenase 1	5351	1p36.22
201021_s_at	0.4644	*DSTN*	destrin (actin depolymerizing factor)	11034	20p12.1
201022_s_at	0.434	*DSTN*	destrin (actin depolymerizing factor)	11034	20p12.1
201032_at	0.4325	*BLCAP*	bladder cancer associated protein	10904	20q11.2-q12
201038_s_at	−0.4044	*ANP32A*	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	8125	15q22.3-q23
201051_at	−0.4105	*ANP32A*	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	8125	15q22.3-q23
201148_s_at	0.4094	*TIMP3*	TIMP metallopeptidase inhibitor 3	7078	22q12.1-q13.2|22q12.3
201204_s_at	0.4182	*RRBP1*	ribosome binding protein 1 homolog 180 kDa (dog)	6238	20p12
201305_x_at	−0.4189	*ANP32B*	acidic (leucine-rich) nuclear phosphoprotein 32 family, member B	10541	9q22.32
201481_s_at	0.405	*PYGB*	phosphorylase, glycogen; brain	5834	20p11.2-p11.1
201500_s_at	0.4222	*PPP1R11*	protein phosphatase 1, regulatory (inhibitor) subunit 11	6992	6p21.3
201525_at	0.475	*APOD*	apolipoprotein D	347	3q26.2-qter
201618_x_at	0.4073	*GPAA1*	glycosylphosphatidylinositol anchor attachment protein 1 homolog (yeast)	8733	8q24.3
201629_s_at	−0.4001	*ACP1*	acid phosphatase 1, soluble	52	2p25
201720_s_at	−0.4154	*LAPTM5*	lysosomal protein transmembrane 5	7805	1p34
201775_s_at	0.4114	*KIAA0494*	KIAA0494	9813	1pter-p22.1
201987_at	−0.4256	*MED13*	mediator complex subunit 13	9969	17q22-q23
202027_at	0.4175	*TMEM184B*	transmembrane protein 184B	25829	22q12
202058_s_at	0.4398	*KPNA1*	karyopherin α 1 (importin α 5)	3836	3q21
202066_at	0.4728	*PPFIA1*	PTPRF interacting protein, α 1	8500	11q13.3
202219_at	0.4223	*SLC6A8*	solute carrier family 6, member 8	6535	Xq28
202421_at	0.4328	*IGSF3*	immunoglobulin superfamily, member 3	3321	1p13
202478_at	0.4103	*TRIB2*	tribbles homolog 2 (*Drosophila*)	28951	2p24.3
202479_s_at	0.5655	*TRIB2*	tribbles homolog 2 (*Drosophila*)	28951	2p24.3
202503_s_at	−0.4017	*KIAA0101*	KIAA0101	9768	15q22.31
202609_at	0.4414	*EPS8*	epidermal growth factor receptor pathway substrate 8	2059	12q13
202629_at	−0.457	*APPBP2*	amyloid β-precursor protein (cytoplasmic tail) binding protein 2	10513	17q21-q23
202818_s_at	0.4258	*TCEB3*	transcription elongation factor B (SIII), polypeptide 3 (110 kDa, elongin A)	6924	1p36.1
202821_s_at	0.4163	*LPP*	LIM domain containing preferred translocation partner in lipoma	4026	3q28
202840_at	−0.5074	*TAF15*	TAF15 RNA polymerase II, TATA box binding protein (TBP)–associated factor	8148	17q11.1-q11.2
202853_s_at	0.4478	*RYK*	RYK receptor-like tyrosine kinase	6259	3q22
202894_at	0.4577	*EPHB4*	EPH receptor B4	2050	7q22
202943_s_at	0.4066	*NAGA*	N-acetylgalactosaminidase, α-	4668	22q13-qter|22q11
202946_s_at	0.4029	*BTBD3*	BTB (POZ) domain containing 3	22903	20p12.2
203054_s_at	0.5192	*TCTA*	T-cell leukemia translocation altered gene	6988	3p21
203137_at	−0.4221	*WTAP*	Wilms tumor 1 associated protein	9589	6q25-q27
203304_at	0.4496	*BAMBI*	BMP and activin membrane–bound inhibitor homolog (*Xenopus laevis*)	25805	10p12.3-p11.2
203455_s_at	0.4052	*SAT1*	spermidine/spermine N1-acetyltransferase 1	6303	Xp22.1
203488_at	0.4243	*LPHN1*	latrophilin 1	22859	19p13.2
203657_s_at	0.4297	*CTSF*	cathepsin F	8722	11q13
203832_at	−0.4357	*SNRPF*	small nuclear ribonucleoprotein polypeptide F	6636	12q23.1
204209_at	0.4157	*PCYT1A*	phosphate cytidylyltransferase 1, choline, α	5130	3q29
204233_s_at	0.4287	*CHKA*	choline kinase α	1119	11q13.2
204266_s_at	0.4633	*CHKA*	choline kinase α	1119	11q13.2
204301_at	−0.4881	*KBTBD11*	kelch repeat and BTB (POZ) domain containing 11	9920	8p23.3
204501_at	0.4265	*NOV*	nephroblastoma overexpressed gene	4856	8q24.1
204528_s_at	−0.4544	*NAP1L1*	nucleosome assembly protein 1-like 1	4673	12q21.2
204542_at	0.4563	*ST6GALNAC2*	ST6 -N-acetylgalactosaminide α-2,6-sialyltransferase 2	10610	17q25.1
204566_at	−0.4076	*PPM1D*	protein phosphatase 1D magnesium–dependent, delta isoform	8493	17q23.2
204626_s_at	0.4075	*ITGB3*	integrin, β_3 (platelet glycoprotein IIIa, antigen CD61)	3690	17q21.32
204638_at	0.4635	*ACP5*	acid phosphatase 5, tartrate resistant	54	19p13.3-p13.2
204653_at	0.4107	*TFAP2A*	transcription factor AP-2 α	7020	6p24
204654_s_at	0.4418	*TFAP2A*	transcription factor AP-2 α	7020	6p24
204783_at	0.4152	*MLF1*	myeloid leukemia factor 1	4291	3q25.1
204789_at	−0.5508	*FMNL1*	formin-like 1	752	17q21
204803_s_at	−0.4276	*RRAD*	Ras-related associated with diabetes	6236	16q22
205209_at	−0.4004	*ACVR1B*	activin A receptor, type IB	91	12q13
205334_at	0.4318	*S100A1*	S100 calcium binding protein A1	6271	1q21
205376_at	0.4913	*INPP4B*	inositol polyphosphate-4-phosphatase, type II, 105 kDa	8821	4q31.21
205441_at	0.4227	*OCEL1*	occludin/ELL domain containing 1	79629	19p13.11
205749_at	0.4594	*CYP1A1*	cytochrome P450, family 1, subfamily A, polypeptide 1	1543	15q24.1
206066_s_at	−0.4726	*RAD51C*	RAD51 homolog C (S. cerevisiae)	5889	17q22-q23
206077_at	0.4308	*KEL*	Kell blood group, metalloendopeptidase	3792	7q33
206305_s_at	−0.4283	*C8A*	complement component 8, α polypeptide	731	1p32
206523_at	0.4377	*CYTH3*	cytohesin 3	9265	7p22.1
206775_at	0.4011	*CUBN*	cubilin (intrinsic factor-cobalamin receptor)	8029	10p12.31
206994_at	0.4623	*CST4*	cystatin S	1472	20p11.21
207039_at	0.4992	*CDKN2A*	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	1029	9p21
207098_s_at	0.419	*MFN1*	mitofusin 1	55669	3q26.33
207714_s_at	0.4095	*SERPINH1*	serpin peptidase inhibitor, clade H, member 1	871	11q13.5
207833_s_at	0.4588	*HLCS*	holocarboxylase synthetase	3141	21q22.1|21q22.13

Table W4. (*continued*)

Probe Set*	Rho†	Symbol	Gene Title	Entrez Gene ID	Cytoband
207913_at	0.426	*CYP2F1*	cytochrome P450, family 2, subfamily F, polypeptide 1	1572	19q13.2
207938_at	0.4719	*PI15*	peptidase inhibitor 15	51050	8q21.11
207992_s_at	−0.407	*AMPD3*	adenosine monophosphate deaminase (isoform E)	272	11p15
208291_s_at	−0.4216	*TH*	tyrosine hydroxylase	7054	11p15.5
208356_x_at	−0.4035	*CSH1*	chorionic somatomammotropin hormone 1 (placental lactogen)	1442	17q24.2
208754_s_at	−0.455	*NAP1L1*	nucleosome assembly protein 1–like 1	4673	12q21.2
208776_at	−0.4355	*PSMD11*	proteasome (prosome, macropain) 26S subunit, non-ATPase, 11	5717	17q11.2
208949_s_at	0.5056	*LGALS3*	lectin, galactoside-binding, soluble, 3	3958	14q21-q22
208978_at	0.4147	*CRIP2*	cysteine-rich protein 2	1397	14q32.3
209054_s_at	−0.4028	*WHSC1*	Wolf-Hirschhorn syndrome candidate 1	7468	4p16.3
209280_at	−0.4334	*MRC2*	mannose receptor, C type 2	9902	17q23.2
209361_s_at	0.4418	*PCBP4*	poly(rC) binding protein 4	57060	3p21
209399_at	0.4037	*HLCS*	holocarboxylase synthetase	3141	21q22.1|21q22.13
209484_s_at	−0.4026	*NSL1*	NSL1, MIND kinetochore complex component, homolog (S. cerevisiae)	25936	1q41
209485_s_at	0.4055	*OSBPL1A*	oxysterol binding protein-like 1A	114876	18q11.1
209598_at	−0.4092	*PNMA2*	paraneoplastic antigen MA2	10687	8p21.2
209736_at	0.4124	*SOX13*	SRY (sex determining region Y)-box 13	9580	1q32
209744_x_at	0.4511	*ITCH*	itchy E3 ubiquitin protein ligase homolog (mouse)	83737	20q11.22-q11.23
209849_s_at	−0.4522	*RAD51C*	RAD51 homolog C (S. cerevisiae)	5889	17q22-q23
209945_s_at	0.4391	*GSK3B*	glycogen synthase kinase 3β	2932	3q13.3
210083_at	0.4181	*SEMA7A*	semaphorin 7A, GPI membrane anchor	8482	15q22.3-q23
210194_at	−0.5138	*PLA2R1*	phospholipase A2 receptor 1, 180 kDa	22925	2q23-q24
210235_s_at	0.4315	*PPFIA1*	PTPRF interacting protein, α 1	8500	11q13.3
210285_x_at	−0.4267	*WTAP*	Wilms tumor 1 associated protein	9589	6q25-q27
210740_s_at	0.411	*ITPK1*	inositol 1,3,4-triphosphate 5/6 kinase	3705	14q31
210854_x_at	0.4757	*SLC6A8*	solute carrier family 6, member 8	6535	Xq28
210980_s_at	0.4375	*ASAH1*	N-acylsphingosine amidohydrolase 1	427	8p22-p21.3
211381_x_at	0.403	*SPAG11B*	sperm associated antigen 11B	10407	8p23-p22
211870_s_at	0.4171	*PCDHA3*	protocadherin α 3	56145	5q31
211883_x_at*	0.4335	*CEACAM1*	carcinoembryonic antigen–related cell adhesion molecule 1	634	19q13.2
211988_at	−0.4108	*SMARCE1*	SWI/SNF related, matrix associated, subfamily e, member 1	6605	17q21.2
212007_at	−0.4351	*UBXN4*	UBX domain protein 4	23190	2q21.3
212056_at	−0.4018	*KIAA0182*	KIAA0182	23199	16q24.1
212086_x_at	0.4143	*LMNA*	lamin A/C	4000	1q21.2-q21.3
212089_at	0.4162	*LMNA*	lamin A/C	4000	1q21.2-q21.3
212226_s_at	0.4283	*PPAP2B*	phosphatidic acid phosphatase type 2B	8613	1pter-p22.1
212230_at	0.4355	*PPAP2B*	phosphatidic acid phosphatase type 2B	8613	1pter-p22.1
212252_at	−0.4583	*CAMKK2*	calcium/calmodulin–dependent protein kinase kinase 2, β	10645	12q24.2
212338_at	0.41	*MYO1D*	myosin ID	4642	17q11-q12
212375_at	−0.4327	*EP400*	E1A binding protein p400	57634	12q24.33
212631_at	0.4084	*STX7*	syntaxin 7	8417	6q23.1
212747_at	0.4647	*ANKS1A*	ankyrin repeat and sterile α motif domain containing 1A	23294	6p21.31
212807_s_at	0.5102	*SORT1*	sortilin 1	6272	1p21.3-p13.1|1p21.3-p13.1
212876_at	0.4171	*B4GALT4*	UDP-Gal:βGlcNAc β 1,4- galactosyltransferase, polypeptide 4	8702	3q13.3
212957_s_at	0.4444	*LOC92249*	hypothetical LOC92249	92249	Xq11.1
213236_at	0.4565	*SASH1*	SAM and SH3 domain containing 1	23328	6q24.3
213242_x_at	0.4868	*KIAA0284*	KIAA0284	283638	14q32.33
213293_s_at	−0.4534	*TRIM22*	tripartite motif–containing 22	10346	11p15
213294_at	−0.4017	—	—	—	—
213310_at	0.4496	*EIF2C2*	Eukaryotic translation initiation factor 2C, 2	27161	8q24
213343_s_at	0.604	*GDPD5*	glycerophosphodiester phosphodiesterase domain containing 5	81544	11q13.4-q13.5
213472_at	−0.4171	*HNRNPH1*	heterogeneous nuclear ribonucleoprotein H1 (H)	3187	5q35.3
213575_at	−0.4166	*TRA2A*	transformer 2 α homolog (*Drosophila*)	29896	7p15.3
213649_at	−0.4247	*SFRS7*	splicing factor, arginine/serine–rich 7, 35 kDa	6432	2p22.1
213672_at	−0.4654	*MARS*	methionyl-tRNA synthetase	4141	12q13.2
213702_x_at	0.4707	*ASAH1*	N-acylsphingosine amidohydrolase 1	427	8p22-p21.3
213843_x_at	0.4967	*SLC6A8*	solute carrier family 6, member 8	6535	Xq28
213864_s_at	−0.4207	*NAP1L1*	nucleosome assembly protein 1–like 1	4673	12q21.2
213902_at	0.4116	*ASAH1*	N-acylsphingosine amidohydrolase 1	427	8p22-p21.3
213921_at	0.4172	*SST*	somatostatin	6750	3q28
213954_at	−0.4715	*FAM169A*	family with sequence similarity 169, member A	26049	5q13.3
214152_at	0.4033	*CCPG1*	cell cycle progression 1	9236	15q21.1
214172_x_at	0.455	*RYK*	RYK receptor–like tyrosine kinase	6259	3q22
214180_at	0.4349	*MAN1C1*	mannosidase, α, class 1C, member 1	57134	1p35
214213_x_at	0.428	*LMNA*	Lamin A/C	4000	1q21.2-q21.3
214280_x_at	−0.4075	*HNRNPA1*	heterogeneous nuclear ribonucleoprotein A1	3178	12q13.1
214584_x_at	0.444	*ACACB*	acetyl-Coenzyme A carboxylase β	32	12q24.11
214635_at	0.4365	*CLDN9*	claudin 9	9080	16p13.3
214812_s_at	−0.5214	*MOBKL1B*	MOB1, Mps One Binder kinase activator–like 1B (yeast)	55233	2p13.1
214971_s_at	0.4022	*ST6GAL1*	ST6 β-galactosamide α-2,6-sialyltranferase 1	6480	3q27-q28
215017_s_at	0.4167	*FNBP1L*	formin binding protein 1-like	54874	1p22.1
215096_s_at	−0.4088	*ESD*	esterase D/formylglutathione hydrolase	2098	13q14.1-q14.2
215381_at	0.4365	*FRAP1*	FK506 binding protein 12-rapamycin associated protein 1	2475	1p36.2
215495_s_at	0.4115	*SAMD4A*	sterile α motif domain containing 4A	23034	14q22.2

Table W4. (*continued*)

Probe Set*	Rho†	Symbol	Gene Title	Entrez Gene ID	Cytoband
215535_s_at	0.4043	*AGPAT1*	1-acylglycerol-3-phosphate O-acyltransferase 1	10554	6p21.3
215611_at	−0.4135	*TCF12*	transcription factor 12	6938	15q21
215693_x_at	0.4599	*DDX27*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	55661	20q13.13
215731_s_at*	−0.4896	*MPHOSPH9*	M-phase phosphoprotein 9	10198	12q24.31
215749_s_at	0.4258	*GORASP1*	golgi reassembly stacking protein 1, 65 kDa	64689	3p22-p21.33
215812_s_at	0.5162	*LOC653562*	similar to solute carrier family 6 member 8	—	16p11.2 /// Xq28
216032_s_at	0.4316	*ERGIC3*	ERGIC and golgi 3	51614	20pter-q12
216060_s_at	0.4403	*DAAM1*	dishevelled associated activator of morphogenesis 1	23002	14q23.1
216086_at	0.4003	*SV2C*	synaptic vesicle glycoprotein 2C	22987	5q13.3
216560_x_at	0.4343	*IGL@*	immunoglobulin lambda locus	3535	22q11.1-q11.2
216629_at	0.4066	*SRRM2*	Serine/arginine repetitive matrix 2	23524	16p13.3
216751_at	−0.4111	—	—	—	—
216835_s_at	−0.4124	*DOK1*	docking protein 1, 62 kDa (downstream of tyrosine kinase 1)	1796	2p13
216874_at	0.489	*DKFZp686O1327*	Hypothetical gene supported by BC043549; BX648102	401014	2q22.3
216976_s_at	0.4007	*RYK*	RYK receptor–like tyrosine kinase	6259	3q22
217036_at	0.4367	*LOC100293679*	hypothetical protein LOC100293679	100293679	—
217094_s_at	0.4157	*ITCH*	itchy E3 ubiquitin protein ligase homolog (mouse)	83737	20q11.22–q11.23
217573_at	0.4175	*GRIN2C*	glutamate receptor, ionotropic, N–methyl D–aspartate 2C	2905	17q25
217606_at	0.4886	—	—	—	—
217613_at	−0.4542	*TMEM144*	transmembrane protein 144	55314	4q32.1
217730_at*	0.4432	*TMBIM1*	transmembrane BAX inhibitor motif containing 1	64114	2p24.3–p24.1
217749_at	0.4832	*COPG*	coatomer protein complex, subunit γ	22820	3q21.3
217925_s_at	0.4097	*C6orf106*	chromosome 6 open reading frame 106	64771	6p21.31
218099_at	0.4599	*TEX2*	testis expressed 2	55852	17q23.3
218299_at	0.4424	*C11orf24*	chromosome 11 open reading frame 24	53838	11q13
218509_at	0.5167	*LPPR2*	lipid phosphate phosphatase–related protein type 2	64748	19p13.2
218670_at	−0.4554	*PUS1*	pseudouridylate synthase 1	80324	12q24.33
218779_x_at	0.4243	*EPS8L1*	EPS8-like 1	54869	19q13.42
218936_s_at	−0.4272	*CCDC59*	coiled-coil domain containing 59	29080	12q21.31
218963_s_at	0.4383	*KRT23*	keratin 23 (histone deacetylase-inducible)	25984	17q21.2
218970_s_at	−0.4909	*CUTC*	cutC copper transporter homolog (*E. coli*)	51076	10q24.2
219011_at	0.4364	*PLEKHA4*	pleckstrin homology domain containing, family A member 4	57664	19q13.33
219046_s_at	0.4165	*PKNOX2*	PBX/knotted 1 homeobox 2	63876	—
219108_x_at	0.4345	*DDX27 /// SS18*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	—	18q11.2 /// 20q13.13
219112_at	−0.4589	*RAPGEF6*	Rap guanine nucleotide exchange factor (GEF) 6	51735	5q31.1
219143_s_at	0.4134	*RPP25*	ribonuclease P/MRP 25 kDa subunit	54913	15q24.2
219278_at	0.4464	*MAP3K6*	mitogen-activated protein kinase kinase kinase 6	9064	1p36.11
219394_at	−0.4245	*PGS1*	phosphatidylglycerophosphate synthase 1	9489	17q25.3
219428_s_at	0.4054	*PXMP4*	peroxisomal membrane protein 4, 24 kDa	11264	20q11.22
219450_at	0.4348	*C4orf19*	chromosome 4 open reading frame 19	55286	4p14
219569_s_at	−0.4059	*TMEM22*	transmembrane protein 22	80723	3q22.3
219710_at	0.4466	*SH3TC2*	SH3 domain and tetratricopeptide repeats 2	79628	5q32
219873_at	0.4197	*COLEC11*	collectin subfamily member 11	78989	2p25.3
220038_at	0.4149	*C8orf44 /// SGK3*	chromosome 8 open reading frame 44	—	8q12.3-q13.1 /// 8q13.1
220499_at	0.443	*FNDC8*	fibronectin type III domain containing 8	54752	17q12
220559_at	0.4341	*EN1*	engrailed homeobox 1	2019	2q13-q21
220948_s_at	0.447	*ATP1A1*	ATPase, Na^+/K^+ transporting, α 1 polypeptide	476	1p21
220999_s_at	−0.4167	*CYFIP2*	cytoplasmic FMR1 interacting protein 2	26999	5q33.3
221017_s_at	0.409	*LRRC3*	leucine rich repeat containing 3	81543	21q22.3
221215_s_at	0.4981	*RIPK4*	receptor-interacting serine-threonine kinase 4	54101	21q22.3
221410_x_at	0.4631	*PCDHB3*	protocadherin β3	56132	5q31
221489_s_at	0.4378	*SPRY4*	sprouty homolog 4 (*Drosophila*)	81848	5q31.3
221616_s_at	−0.4113	*TAF9B*	TAF9B RNA polymerase II	51616	Xq13.1-q21.1
221681_s_at	−0.4101	*DSPP*	dentin sialophosphoprotein	1834	4q21.3
221683_s_at	−0.4697	*CEP290*	centrosomal protein 290 kDa	80184	12q21.32
221738_at	0.4105	*KIAA1219*	KIAA1219	57148	20q11.23
221816_s_at	−0.4902	*PHF11*	PHD finger protein 11	51131	13q14.2
221819_at	−0.4578	*RAB35*	RAB35, member RAS oncogene family	11021	12q24.31
221821_s_at*	−0.4137	*C12orf41*	chromosome 12 open reading frame 41	54934	12q13.11
221900_at	−0.4483	*COL8A2*	collagen, type VIII, α 2	1296	1p34.2
221919_at	−0.4026	*HNRNPA1*	heterogeneous nuclear ribonucleoprotein A1	3178	12q13.1
221965_at*	−0.5905	*MPHOSPH9*	M-phase phosphoprotein 9	10198	12q24.31
221986_s_at	0.4164	*KLHL24*	kelch-like 24 (*Drosophila*)	54800	3q27.1
222347_at	0.4183	*LOC644450*	hypothetical protein LOC644450	644450	1q12
32502_at	0.4446	*GDPD5*	glycerophosphodiester phosphodiesterase domain containing 5	81544	11q13.4-q13.5
36888_at	−0.4388	*HAUS5*	HAUS augmin–like complex, subunit 5	23354	19q13.12
38918_at	0.4097	*SOX13*	SRY (sex determining region Y)-box 13	9580	1q32
41644_at	0.4485	*SASH1*	SAM and SH3 domain containing 1	23328	6q24.3
52164_at	0.4047	*C11orf24*	chromosome 11 open reading frame 24	53838	11q13
52169_at	−0.4345	*STRADA*	STE20-related kinase adaptor α	92335	17q23.3
64899_at	0.4262	*LPPR2*	lipid phosphate phosphatase–related protein type 2	64748	19p13.2

*Probe set designation from the Affymetrix HG-U133A platform. The five probes exhibiting concordant expression across all three data sets are specifically asterisked.

†Correlation coefficient (ρ) for Spearman rank–based correlation of indicated probe expression in the NCI-60 cell line data set, *GSE5720*, to the C1311 IC_{50} values for each of the sixty cell lines. A cutoff of 0.4 was used for biomarker discovery.