Prediction of drug combination chemosensitivity in human bladder cancer

Dmytro M. Havaleshko,[1] HyungJun Cho,[2,4] Mark Conaway,[2] Charles R. Owens,[1] Garret Hampton,[5] Jae K. Lee,[2] and Dan Theodorescu[1,3]

Departments of [1]Molecular Physiology and [2]Public Health Sciences and [3]Mellon Prostate Cancer Institute, University of Virginia, Charlottesville, Virginia; [4]Department of Statistics, Korea University, Seoul, Korea; and [5]Genomics Institute of the Novartis Research Foundation, San Diego, California

Abstract

The choice of therapy for metastatic cancer is largely empirical because of a lack of chemosensitivity prediction for available combination chemotherapeutic regimens. Here, we identify molecular models of bladder carcinoma chemosensitivity based on gene expression for three widely used chemotherapeutic agents: cisplatin, paclitaxel, and gemcitabine. We measured the growth inhibition elicited by these three agents in a series of 40 human urothelial cancer cell lines and correlated the GI_{50} (50% of growth inhibition) values with quantitative measures of global gene expression to derive models of chemosensitivity using a misclassification-penalized posterior approach. The misclassification-penalized posterior–derived models predicted the growth response of human bladder cancer cell lines to each of the three agents with sensitivities of between 0.93 and 0.96. We then developed an *in silico* approach to predict the cellular growth responses for each of these agents in the clinically relevant two-agent combinations. These predictions were prospectively evaluated on a series of 15 randomly chosen bladder carcinoma cell lines. Overall, 80% of the predicted combinations were correct ($P = 0.0002$). Together, our results suggest that chemosensitivity to drug combinations can be predicted based on molecular models and provide the framework for evaluation of such models in patients undergoing combination chemotherapy for cancer. If validated *in vivo*, such predictive models have the potential to guide therapeutic choice at the level of an individual's tumor. [Mol Cancer Ther 2007;6(2):578–86]

Introduction

Most patients with epithelial cancers requiring systemic treatment undergo combination chemotherapy. However, a major challenge in these patients has been the prediction of chemotherapeutic efficacy based on the biological profile of the tumor. Recent studies have indicated that chemotherapeutic sensitivity of the NCI-60 cancer cell line panel can be, in part, predicted by gene expression (1). However, the predictive utility of these molecular correlates of chemosensitivity has not yet been validated for other cellular histologies or for drug combinations. Both issues have been partially addressed in bladder cancer where tumor expression profiling has been used to predict responsiveness to one combination chemotherapy (2). Unfortunately, the practical consequences of this result are limited because this combination is infrequently used, and similar studies are laking for more popular combinations. Furthermore, with the constant arrival of new agents, there is a steady evolution of combination chemotherapy regimens. Together, these two facts complicate the selection process of the best combination therapy for individual patients.

Here, we address this limitation by developing a novel approach that predicts chemotherapeutic responsiveness of human bladder cancer cells to combinations of gemcitabine with cisplatin (3), cisplatin with paclitaxel (4), and gemcitabine with paclitaxel (5) that have shown promising results in clinical trials. Furthermore, by virtue of using single-drug sensitivities to mathematically predict combination effects (rather than using effects of combination directly), this approach has the unique advantage of allowing the evaluation of any number of agents in combination and of allowing the integration of new agents into new combinations as needed.

Materials and Methods

Cell Lines, Cell Culture, Gene Expression Profiling, and Dose-Response Data Generation and Analyses

The human bladder cancer cell lines and the respective growth conditions used in this study have been previously described (6, 7). Cisplatin was purchased from Sigma (St. Louis, MO), dissolved in Dulbecco's PBS, and aliquoted in 1 mg/mL stocks. Paclitaxel was purchased from Sigma, dissolved in DMSO, and aliquoted in 1 mmol/L stocks. Gemcitabine was purchased from the University of Virginia Medical Center Pharmacy, dissolved in PBS, and aliquoted in 0.1 mol/L stocks. Cell lines were maintained in appropriate media, in a humidified atmosphere containing

Received 8/15/06; revised 10/28/06; accepted 12/18/06.

Grant support: NIH grant CA075115.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Dan Theodorescu, Department of Molecular Physiology and Biological Physics, University of Virginia Health Sciences Center, Box 422, Charlottesville, VA 22908. Phone: 434-924-0042; Fax: 434-982-3652. E-mail: dt9d@virginia.edu

doi:10.1158/1535-7163.MCT-06-0497

5% CO_2 in air, except CRL2169 (SW780), which requires no CO_2 for its growth. Cell lines were subcultured in an aqueous solution of 0.05% trypsin (GibcoBRL, Grand Island, NY; 1:250) and 0.016% EDTA. Each cell line was used within 10 passages from its archival passage number to minimize any long-term cell culture effects. Gene expression analysis of bladder cell lines was carried out as previously described using the HG-U133A GeneChip array (Affymetrix, Santa Clara, CA; refs. 6, 7). The image file was analyzed with robust microarray analysis to obtain the expression intensity values of the microarray data (8).

Cell lines were seeded in 96-well cell culture plates (Costar, Corning, NY) at a density of 1,000 per well. Twenty-four hours later, cells were exposed to the drugs diluted in RPMI 1640, containing 10% fetal bovine serum, a concentration that is required by >75% of cell lines for their normal growth, at a total volume of 200 μL. Each drug dose was plated in triplicate, and the experiment was repeated four to seven times. The doses for cisplatin were 200, 400, 800, 1,600, 3,200, and 6,400 ng/mL; for paclitaxel were 0.0001, 0.001, 0.002, 0.005, 0.01, and 0.1 μmol/L; for gemcitabine were 0.001, 0.01, 0.1, 1, 10, and 100 μmol/L. Plates were incubated for 72 h with carrier or drug, and growth inhibition was assessed by Alamar Blue (BioSource International, Inc., Camarillo, CA; refs. 9, 10). Our doses for cisplatin, paclitaxel, and gemcitabine were chosen to be similar to the range of doses used by the National Cancer Institute (NCI) in their screening of the NCI-60 set of cell lines.[6]

Estimation of GI_{50} Values

From the dose-response data, $\log_{10}(GI_{50})$ values (log base 10 of concentration required to inhibit cell growth by 50% in comparison with untreated control) were estimated for all the cell lines by deriving log(dose) concentration curves on cell count percents as described below. To estimate the GI_{50} values reliably, we computed Euclidean distances among all replicated experiments and excluded outlying experiments if they were in the top 20% among all measured distances. This percentage was determined heuristically based on the general observations in experimental quality control. Furthermore, we did not see significant changes in our results by slightly changing this proportion as several replicated experiments were averaged to estimate our GI_{50} values (data not shown). Subsequently, the data were fitted to the following nonlinear regression model for estimating dose-response curve of each cell line:

$$\text{Percent} = 1 - 1/(1 + \exp(-(\log_{10}(\text{dose}) - \beta)/\alpha),$$

where α and β determine the shape of a fitted line. This sigmoidal regression function was used to capture the natural shapes of drug dose responses. Thus, the estimated β is the predicted $\log_{10}(GI_{50})$ value, the expected log concentration achieving the cell count reduction of 50%. Similarly, $\log_{10}(GI_{30})$ and $\log_{10}(GI_{70})$ values (i.e., the concentrations required to inhibit cell growth by 30% and 70%, respectively, in comparison with untreated control) were also calculated.

Determination of Sensitive and Resistant Cell Lines for Single Drug Sensitivity

Cell line drug sensitivity was classified using the GI estimates and application of a criterion dose concept. We defined the criterion dose as the minimum $\log_{10}$(drug dose) among the experimental dose concentrations of each compound at which at least 25% of the cell lines showed growth inhibition >50%. Criterion doses were determined as $\log_{10}$(400 ng/mL) for cisplatin, $\log_{10}$(0.005 μmol/L) for paclitaxel, and $\log_{10}$(0.1 μmol/L) for gemcitabine, which provided at least 10 drug-''sensitive'' cell lines for each drug. Using these criterion dose concentrations, each cell line was defined as sensitive if $\log_{10}(GI_{50}) \leq$ criterion dose; strongly sensitive if $\log_{10}(GI_{30}) <$ criterion dose, or resistant if $\log_{10}(GI_{70}) >$ criterion dose; and intermediate if $\log_{10}(GI_{50}) >$ criterion dose and $\log_{10}(GI_{70}) <$ criterion dose.

Statistical Discovery of Molecular Chemosensitivity Prediction Models for Single Drugs

For statistical discovery of prediction models, all 22,215 genes on the HG-U133A array were first evaluated for their ability to differentiate sensitive and resistant cell lines; intermediate lines were excluded from the analysis. The most significant genes were selected both by local pooled error test (11) and significance analysis of microarray method (12). After candidate biomarker probes were identified for each tested compound on the basis of significant differential expression for drug sensitivity, we next searched among those candidate biomarkers for ones that would form optimal parsimonious models for prediction of the activity of the compound. For this, we used the ''misclassification-penalized posterior'' (MiPP) algorithm, which we introduced previously and is available at the open-source Bioconductor web site[7] (13). MiPP is based on stepwise incremental classification modeling discovery for the optimal, most parsimonious prediction models and double cross-validated evaluation for each trained prediction model. Model training can be done from several different classification modeling techniques such as linear discriminant analysis, quadratic discriminant analysis, support vector machines, or logistic regression; linear discriminant analysis was used for most application in our current study. In the double cross-validation, the first cross-validation is based on random splitting of the whole data set into a training set and an independent test set for external model validation. The second is an n-fold cross-validation on the training set to avoid the pitfalls of a large-screening search and to obtain the most parsimonious optimal prediction models. Independent splits of the data result in multiple prediction models. MiPP generates multiple independent splits, which, in turn, results in multiple prediction models. The multiple models from different splits were re-evaluated

[6] http://dtp.nci.nih.gov

[7] www.bioconductor.org

on a large number of (e.g., 100) random splits of test and training sets to obtain their objective confidence bounds with the summary index, the so-called standardized MiPP (sMiPP score), which varies between −1 to 1, from the worst to the best. From this confidence interval evaluation, mean and lower 5% sMiPP scores were obtained for each of the candidate prediction models, together with mean misclassification rates (error rate). The final prediction of sensitive (or resistant) cell lines was done by averaging its (posterior) classification probabilities of the top three prediction models exceeding 5% sMiPP > 0.5. In performing MiPP analysis, we used the default values for many tuning variables of the MiPP Bioconductor R package. For example, n.fold, p.test, n.split, and n.seq were 5, 1/3, 20, and 3, respectively. However, we preselected the most significant top 1% genes by local pooled error and significance analysis of microarray and did not use the MiPP gene selection option by setting percent.cut = 0.

Statistical Chemosensitivity Prediction for Combination Drug Treatments

Prediction of combination drug efficacy was obtained based on the final single-drug prediction models, directly using classification probabilities of each cell line from these models. That is, assuming two different drug compounds acted independently, the combination chemosensitivity probability P_{AB} of their combination treatment was derived as:

$$1 - P_A[\text{resistant for drug A}] \times P_B[\text{resistant for drug B}].$$

Here, P_A and P_B are the chemosensitivity response probabilities based on the prediction models for compound A and B, respectively. Because this provides a somewhat optimistic probability evaluation of chemosensitivity (e.g., if $P_A = P_B = 0.5$, then $P_{AB} = 0.75$), we used a strict decision criterion ($P_{AB} \geq 0.75$) for predicting chemosensitivity of each cell line to combination treatment.

Results

Evaluation of *In vitro* Drug Sensitivity of Human Bladder Cell Lines to Single Agents

To approach the development of molecular models of chemotherapeutic sensitivity in human bladder cancer, we focused on a well-defined series of 40 urothelial cell lines for which we could measure sensitivity to relevant chemotherapeutic agents *in vitro* and correlate these responses with global measurements of gene expression (Fig. 1). The *in vitro* sensitivity of these 40 bladder cancer cell lines to cisplatin, paclitaxel, and gemcitabine was carried out as described in Materials and Methods. Typical dose-response curves for representative sensitive and resistant cell lines are shown for each agent in Fig. 2A. The cell lines were then divided into three groups (sensitive, intermediate, and resistant) based on GI estimates and the criterion dose (defined in Materials and Methods). Figure 2B to D shows the $\log_{10}(GI_{30})$, $\log_{10}(GI_{50})$, and $\log_{10}(GI_{70})$ of the 40 cell lines for each of the agents. For cisplatin, we identified 16 sensitive and 11 resistant cell lines (Fig. 2B), 17 sensitive and 11 resistant cell lines for paclitaxel (Fig. 2C), and 8 sensitive and 11 resistant for gemcitabine (Fig. 2D). Cell lines that did not meet the ''sensitive/resistant'' criteria were excluded from further analyses. For some cell lines, log(GI) values could not be

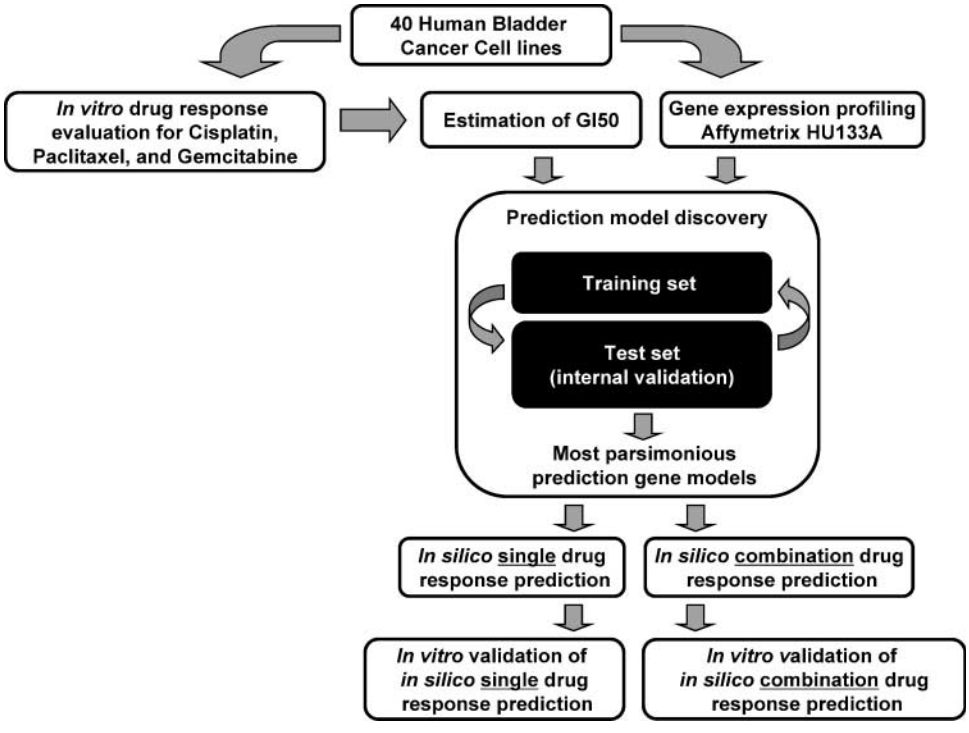

Figure 1. Schematic diagram for the overall process of discovery and validation of gene expression models of chemosensitivity by transcriptional profiling.

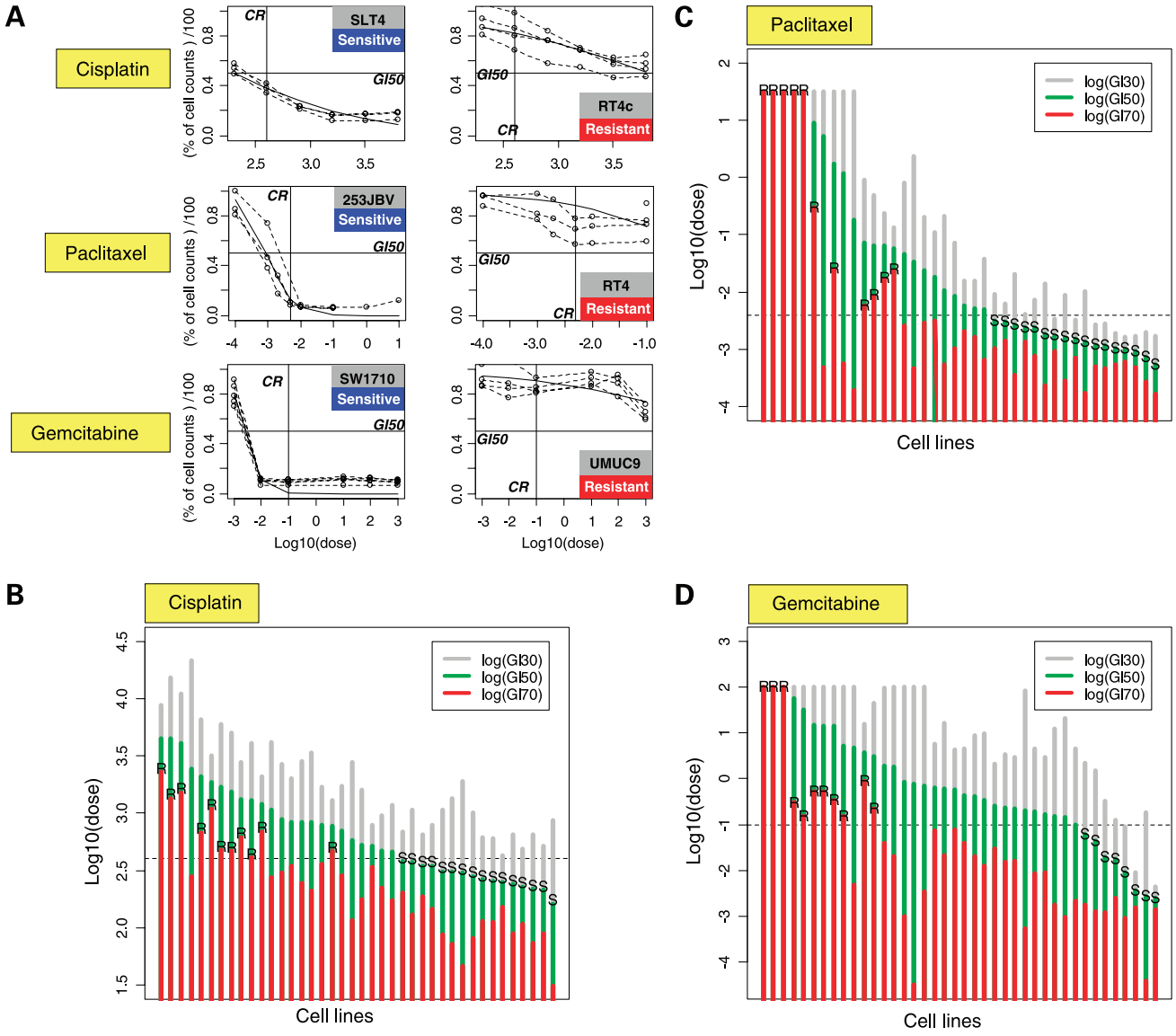

Figure 2. Classification of sensitive and resistant cell lines to single-drug chemotherapy. **A,** growth inhibition dose response curves for SLT4 and RT4 in respond to cisplatin, 253-JBV and RT4 to paclitaxel, and SW1710 and UMUC9 to gemcitabine. The percentage of cell counts (divided by 100) is indicated on the *Y* axis. Cell lines were defined as sensitive if GI_{50} values were below the dose indicated by the vertical criterion line (criterion dose), whereas resistant cell lines had GI_{50} values above this dose. Cisplatin $\log_{10}$(400 ng/mL), paclitaxel $\log_{10}$(0.005 μmol/L), and gemcitabine $\log_{10}$(0.1 μmol/L). Each individual experiment is indicated by a dotted line. The cell lines were ordered by their GI_{50} values. The fitted nonlinear regression line (*solid curve*) represents the combined estimate. Determination of sensitive (*S*) and resistant (*R*) cell lines to (**B**) cisplatin, (**C**) paclitaxel, and (**D**) gemcitabine. *Gray,* $\log_{10}(GI_{30})$; *green,* $\log_{10}(GI_{50})$; and *red,* $\log_{10}(GI_{70})$ of the 40 cell lines.

estimated due to flat response curves in nonlinear regression model fitting; thus, these log(GI) values of cell lines were thresholded at the maximum dose concentration and were classified as resistant.

Prediction Models for Single Drug Sensitivity

We used the MiPP approach to identify models comprised of gene transcript levels that predicted sensitivity to cisplatin, paclitaxel, and gemcitabine (see Materials and Methods). For cisplatin and paclitaxel, we identified three prediction models that met the criteria for selection of sensitive and resistant cells (i.e., with the lower 5% sMiPP > 0.5). For gemcitabine, we identified only one model that met these criteria (Table 1). The selection and order of these models were based on the 5% sMiPP, so was the order of the models. The mean sMiPPs among the three models for cisplatin were 0.820 to 0.858, with mean misclassification rates of 5.4% to 6.9% (prediction accuracy, 93.1–94.6%), based on independent-set cross-validation as described (13). The prediction performance of paclitaxel models was similar to that of cisplatin with mean misclassification rates

of between 4.1% and 7.1% and mean sMiPPs of 0.830 to 0.910. For gemcitabine, we identified a single model with an associated error rate of 9.6% and sMiPP of 0.742. In addition to the performance calculations above, the utility of these gene models in predicting the responsiveness of these drugs can be appreciated by plotting the expression intensities (log 2 scale) of the first two genes in each of our gene prediction models, adding each classification decision line to show the relationship with our classification modeling (Fig. 3A–C).

Prediction Models for Combination Drug Sensitivity

Given the ability to predict single-drug efficacy *in vitro*, we next asked whether this approach could be used to predict the efficacy of the three commonly used drug doublet combinations in the same types of cells. We applied the same basic MiPP approach but averaged the posterior probabilities from each of the models in cases where more than one model met the criterion dose (i.e., for paclitaxel and cisplatin) and then computed the chemosensitivity probability for a given drug. If the combined posterior probability of chemosensitivity for a drug combination was >0.75, a cell line was predicted to be sensitive to that drug combination.

We evaluated the performance of these *in silico* predictions by randomly selecting 15 of the 40 bladder carcinoma cell lines, attempting to roughly balance the numbers of

Table 1. Best gene prediction models for single-drug chemosensitivity response prediction to cisplatin, paclitaxel, and gemcitabine

Models	Probe set ID	Gene symbol	Gene title
Cisplatin			
Model 1			
Mean ER = 0.069, mean sMiPP = 0.858, lower 5% sMiPP = 0.771	212508_at	*MOAP1*	Modulator of apoptosis 1
	218280_x_at	*HIST2H2AA*	Histone 2, H2aa
	222275_at	*MRPS30*	Mitochondrial ribosomal protein S30
	211573_x_at	*TGM2*	Transglutaminase 2
Model 2			
Mean ER = 0.054, mean sMiPP = 0.860, lower 5% sMiPP = 0.730	212508_at	*MOAP1*	Modulator of apoptosis 1
	203323_at	*CAV2*	Caveolin 2
	208885_at	*LCP1*	Lymphocyte cytosolic protein 1 (L-plastin)
Model 3			
Mean ER = 0.066, mean sMiPP = 0.820, lower 5% sMiPP = 0.715	211559_s_at	*CCNG2*	Cyclin G2
	212094_at	*PEG10*	Paternally expressed 10
	221029_s_at	*WNT5B*	Wingless-type MMTV integration site family, member 5B /// wingless-type MMTV integration site family, member 5B
Paclitaxel			
Model 1			
Mean ER = 0.041, mean sMiPP = 0.910, lower 5% sMiPP = 0.788	214858_at	*GPC1*	Glypican 1
	201860_s_at	*PLAT*	Plasminogen activator, tissue
	201317_s_at	*PSMA2*	Proteasome (prosome, macropain) subunit, α type 2
	211812_s_at	*B3GALT3*	UDP-Gal:betaGlcNAc β1,3-galactosyltransferase, polypeptide 3
	204557_s_at	*DZIP1*	DAZ interacting protein 1
Model 2			
Mean ER = 0.051, mean sMiPP = 0.877, lower 5% sMiPP = 0.770	217728_at	*S100A6*	S100 calcium binding protein A6 (calcyclin)
	206364_at	*KIF14*	Kinesin family member 14
	203741_s_at	*ADCY7*	Adenylate cyclase 7
	203438_at	*STC2*	Stanniocalcin 2
	201105_at	*LGALS1*	Lectin, galactoside-binding, soluble, 1 (galectin 1)
Model 3			
Mean ER = 0.071, mean sMiPP = 0.830, lower 5% sMiPP = 0.746	206059_at	*ZNF91*	Zinc finger protein 91 (HPF7, HTF10)
	209310_s_at	*CASP4*	Caspase 4, apoptosis-related cysteine protease
	213849_s_at	*PPP2R2B*	Protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), β isoform
	202591_s_at	*SBP1*	Single-stranded DNA binding protein 1
Gemcitabine			
Model 1			
Mean ER = 0.096, mean sMiPP = 0.742, lower 5% sMiPP = 0.582	202838_at	*FUCA1*	Fucosidase, α-L-1, tissue
	212206_s_at	*H2AFV*	H2A histone family, member V

NOTE: Up to three models were selected with the selection criterion 5% sMiPP > 0.5.
Abbreviation: ER, error rate.

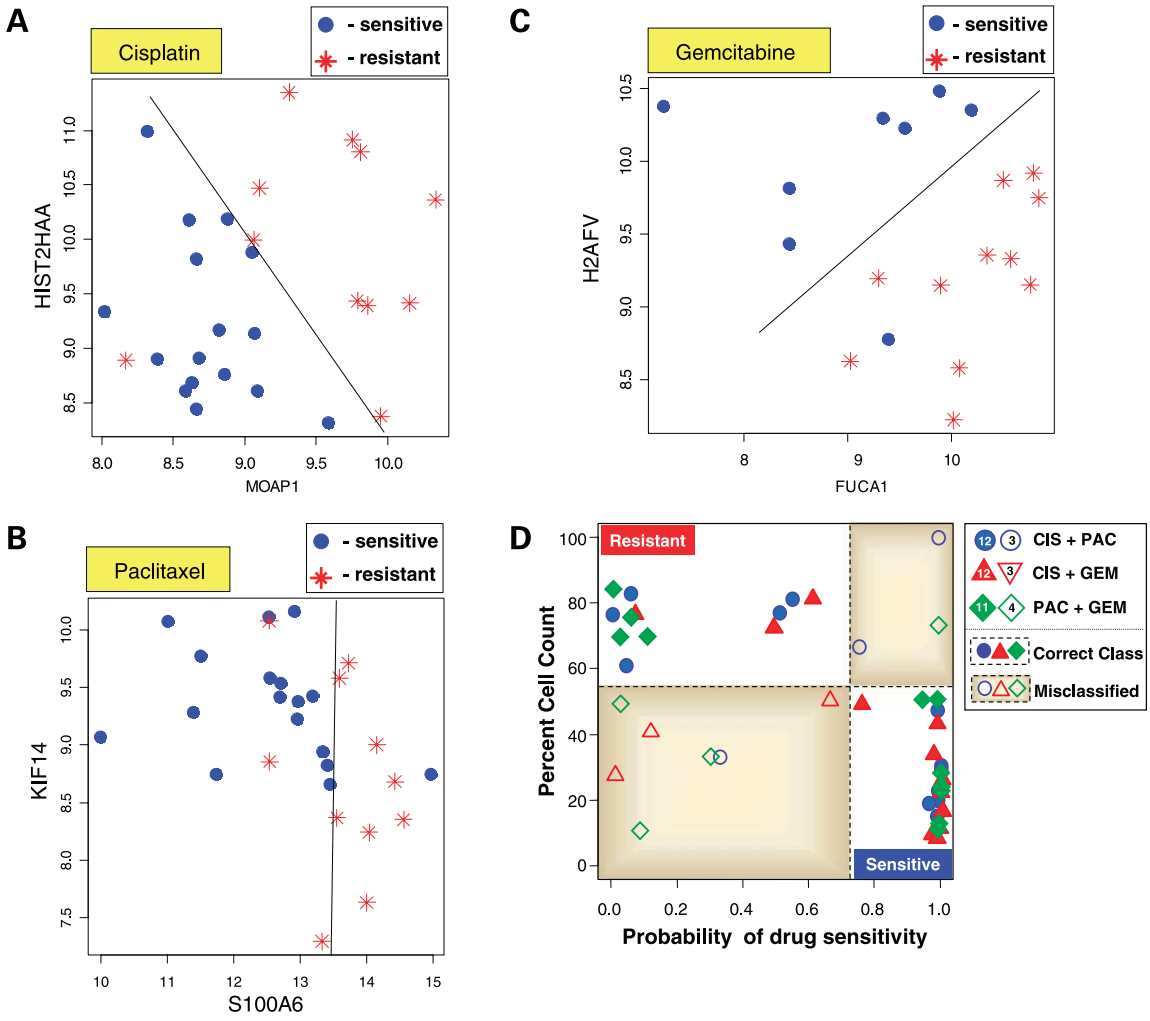

Figure 3. The two-dimensional scatter plots of expression intensities (log 2 scale) of the first two gene prediction models demonstrating their classification performance: (**A**) cisplatin (model 1 in Table 1), (**B**) paclitaxel (model 2 in Table 1), and (**C**) gemcitabine (model 1 in Table 1). Sensitive cells (*blue dots*) and resistant cell lines (*red stars*) were found to be separated by the two selected genes, although some of them were still misclassified. Some of the misclassified ones were better separated by the additional genes; thus, the mean ERs were 0.069, 0.051, and 0.096 for cisplatin, paclitaxel, and gemcitabine, respectively. **D,** the scatter plot of the percent of cell counts compared to control (no drug) versus the posterior probability of sensitivity for the 15 cell lines randomly selected for the evaluation of chemotherapeutic sensitivity prediction for the three two-drug combinations shown. The horizontal (55%) and vertical (0.75) dotted lines divided cell lines into sensitive and resistant based on the percentage of cell count and the posterior probability of sensitivity, respectively. *Cis,* cisplatin; *Pac,* paclitaxel; *Gem,* gemcitabine.

predicted sensitive and resistant cell lines across the three drug combinations. We used the single-drug criteria dose and exposed cells to both drugs simultaneously. The growth of cell lines exposed to the drug combinations compared with control (no drug) was expected to be <55% for sensitive and >55% for resistant cell lines at these doses.

Overall, 35 of the 45 predictions were correct (binomial test, $P = 0.0002$; Table 2; Fig. 3D). Twelve of 15 cell lines (80%; binomial test, $P = 0.03$) were predicted correctly for the cisplatin-paclitaxel combination. Of the three misclassified cell lines, one sensitive line was predicted as resistant, and two resistant cell lines were predicted as sensitive. For the cisplatin-gemcitabine combination, 12 of 15 lines were also predicted correctly; three sensitive cell lines were incorrectly predicted as resistant (80% accuracy; binomial test, $P = 0.03$). Finally, for the combination of paclitaxel and gemcitabine, 11 of 15 lines were correctly classified; three sensitive and one resistant cell lines were misclassified as resistant and sensitive, respectively (73% accuracy; binomial test, $P = 0.11$).

Potential Synergistic Activities with Combination Treatments

In clinical practice, combination treatments significantly outperform single-drug counterparts in treating different types of cancer, either by additive or synergistic drug action. To this end, we found that 7 of 19 (37%) cell lines that were predicted as resistant to the drug combination used were indeed sensitive to the combination when tested,

although the cells were not sensitive to the single compounds of the combination. For example, in the combination treatment of cisplatin and gemcitabine, all three misclassified cases turned out to be predicted resistant cell lines being in fact sensitive when tested. In contrast, fewer (12%, 3 of 26) predicted sensitive cell lines to the drug combination were found to be resistant to the combination (two-sample proportion test, $P = 0.049$).

Discussion

Here, we combined a novel mathematical approach (MiPP probabilities) with comprehensive gene expression profiles of 40 urothelial cell lines to discover high-performance molecular prediction models for single and combination chemotherapeutic sensitivity. The high-performance characteristics of the predictive models obtained in this study may be due to several factors. First, we used a panel of cancer cell lines derived from only one histologic type: urothelial cancer. In contrast to the NCI-60 cancer cell panel, which is comprised of cell lines from multiple anatomic origins, a single anatomic origin should eliminate confounding and biased gene expression signals that represent tissue-dependent sensitivity to different chemotherapy agents. Furthermore, the majority of the cell lines used in this study (14) are derived from invasive or metastatic human urothelial tumors that represent the typical patient population that would receive systemic chemotherapy. Hence, we anticipate that these prediction models may be applicable to clinical urothelial cancer. This conclusion is supported by the observation that cisplatin, a drug used in current clinical treatment of urothelial cancer (15), was highly effective in our assay (i.e., 16 of 40 cell lines meeting the chemosensitive criterion).

To identify gene prediction models for chemosensitivity, we used the MiPP method (13). Several studies have shown good predictive classification of cancer subtypes and prognosis using methods that require large numbers of (>50) genes (16–18), whereas models that are dependent on only a small number of predictive genes has been limited despite the obvious practical advantages. The MiPP method combines the best of both approaches by maintaining excellent predictive accuracy with a small set of genes that are easy to evaluate in human tumors using currently available techniques such as real-time reverse transcription-PCR. This feature is a significant advantage as we begin to prospectively evaluate these genes for their ability to predict tumor response in patients treated with drug combinations.

The approach taken here led to the identification of predictive gene models for each of the three drugs. Cisplatin model 1 is comprised of *TGM2*, *MOAP1*, *HIST2H2AA*, and *MRPS30*; model 2 contains *CAV2*, *LCP1*, and *MOAP1*; and model 3 includes *CCNG2*, *PEG10*, and *WNT5B*. By examining the function of the genes encompassed by these models, a common functional theme was noted, that is, their direct (*TGM2*, *MOAP1*, and *CAV2*) or indirect (*HIST2H2AA* and *LCP1*) participation in apoptosis. Modulator of apoptosis 1 (*MOAP1*) is an important component of the pathway that links death receptors and

Table 2. Predicted sensitivity probabilities to combination therapy and validation in 15 urothelial cancer cell lines

Cell lines	Probability			% Of cell count			Probability			Prediction		
	CIS	PAC	GEM	CIS + PAC	CIS + GEM	PAC + GEM	CIS + PAC	CIS + GEM	PAC + GEM	CIS + PAC	CIS + GEM	PAC + GEM
253JBV	0.90	1.00	0.93	15	25	22	1.00	0.99	1.00	S	S	S
253JLaval	0.54	1.00	0.16	100	81	73	1.00	0.61	1.00	S*	R	S*
253JP	0.98	1.00	0.93	20	16	27	1.00	1.00	1.00	S	S	S
CRL7833	0.97	0.02	0.07	19	10	11	0.97	0.98	0.09	S	S	R*
HT1197	0.50	0.03	0.00	77	72	49	0.51	0.50	0.03	R	R	R*
HT1376	0.54	0.03	0.28	81	50	33	0.55	0.67	0.30	R	R*	R*
HU456	0.95	0.86	0.63	48	34	51	0.99	0.98	0.95	S	S	S
J82	0.33	0.00	0.99	32	43	51	0.33	0.99	0.99	R*	S	S
JON	0.02	0.01	0.10	61	41	70	0.03	0.12	0.11	R	R*	R
MGHU3	0.01	0.01	0.01	76	28	84	0.01	0.01	0.01	R	R*	R
RT4	0.05	0.00	0.02	83	77	70	0.05	0.08	0.02	R	R	R
T24T	1.00	0.99	0.30	29	11	12	1.00	1.00	0.99	S	S	S
TCCSUP	0.75	0.02	0.05	64	49	76	0.76	0.76	0.06	S*	S	R
UMUC3	0.99	0.99	1.00	30	21	21	1.00	1.00	1.00	S	S	S
UMUC6	0.90	0.99	0.91	22	8	13	1.00	0.99	1.00	S	S	S

NOTE: The growth inhibition of the combination drug treatment experiments (% of cell count in cells not exposed to drug) was obtained using the dose concentrations: cisplatin, $\log_{10}$(400 ng/mL); paclitaxel, $\log_{10}$(0.005 μmol/L); and gemcitabine, $\log_{10}$(0.1 μmol/L). A cell line with the larger posterior probability (PP) is more likely to be a sensitive. Single-drug PPs were obtained by averaging posterior probabilities if there were more than one model, and the combined posterior probability is $1 - P_r$ (resistant by cisplatin) $\times P_r$ (resistant by paclitaxel). (predicted as sensitive if PP > 0.75 and as resistant if PP < 0.75). Predicted sensitive or resistant cell lines to the combination pairs of three drug treatments.

Abbreviations: CIS, cisplatin; PAC, paclitaxel; GEM, gemcitabine; S, sensitive; R, resistant.

*Misclassified samples when compared with *in vitro* evaluation of drug combinations.

the apoptotic machinery (19). Caveolin 2 (*CAV2*) is a major component of the inner surface of caveolae and is implicated in the control of cellular growth, signal transduction, lipid metabolism, and apoptosis (20). Lymphocyte cytosolic protein 1 (*LCP1*) is found in hemopoietic cell lineages and also in many types of malignant human cells of non-hemopoietic origin. Cyclin G2 (*CCNG2*) is a member of the cyclin family. Northern blot analysis revealed that cyclin G2 mRNA fluctuates throughout the cell cycle with peak expression in late S phase. Furthermore, cyclin G2 is induced by the DNA-damaging agent actinomycin D (21).

Models for paclitaxel included several genes involved in essential eukaryotic cell functions, such as protein modification (*PLAT*), spermatogenesis and cell differentiation (*DZIP1*), and negative autocrine growth factor regulation (*LGALS1*). However, perhaps the most interesting of this group is *KIF14*. This gene is responsible for microtubule motor activity and is expressed at very low levels in normal tissue samples, compared with significantly increased expression in the majority of tumor samples (22). Its overexpression may lead to rapid mitoses, potentially leading to aneuploidy. *KIF14* overexpression is most striking in retinoblastoma, lung, breast, thymus, and tumors and associated with decreased survival in lung cancer (22). This relationship to paclitaxel sensitivity is intriguing because this drug promotes the assembly of microtubules from tubulin dimers and stabilizes microtubules by preventing depolymerization, thus inducing abnormal arrays of microtubules throughout the cell cycle. We are in the process of investigating this further.

Despite the robust and impressive predictive power of this approach and the lack of viable alternatives, our technique does have limitations. The most significant is its inability to predict drug synergy. For example, a mechanistic basis for gemcitabine/cisplatin (23), paclitaxel/cisplatin (24), and paclitaxel/gemcitabine (25) cytotoxic synergies have been described. Our model was less effective at predicting positive synergy (37%) than negative synergy or antagonism (12%). Fortunately, this differential performance is clinically useful because it implies that drug combinations predicted to be effective in a particular patient using our models are likely to be so and thus can be used as first line treatment. Conversely, those combinations that are predicted to be ineffective may have a low chance of being efficacious based on synergy and thus can be used as ''salvage'' therapy. Furthermore, it is possible that our approach could also be used to develop predictive models for drug synergy. However, this would require the generation of extensive *in vitro* drug response data for each of the different combinations. However, given the biological and clinical relevance of this issue, efforts in this direction are clearly warranted.

Our use of MiPP prediction model was mainly based on the linear discriminate analysis, one of its classification method options in this study. However, other classification options, such as quadratic discriminant analysis, support vector machines, and logistic regression methods, may be further explored in a future study. In addition, our statistical prediction approach to the combination treatments was relatively simple based on the independence assumption of any two compounds. Because our goal is to predict combination responses by two independent experiments of two or multiple compounds, this assumption is statistically inevitable. However, if such drug experiments are done on the same panel of cancer cell lines, we may be able to optimize such a prediction by considering their drug sensitivity through their gene models and then searching for the combination treatments with the highest combined response rates.

In summary, we have developed and validated a novel molecular chemosensitivity prediction model for commonly used combinations of cisplatin, paclitaxel, and gemcitabine, using only the results of their individual drug responses. We believe this prediction strategy warrants prospective validation in the clinical setting and, given the parsimonious nature of the predictions shown here, should be straightforward to implement.

References

1. Scherf U, Ross DT, Waltham M, et al. A gene expression database for the molecular pharmacology of cancer. Nat Genet 2000;24:236 – 44.

2. Takata R, Katagiri T, Kanehira M, et al. Predicting response to methotrexate, vinblastine, doxorubicin, and cisplatin neoadjuvant chemotherapy for bladder cancers through genome-wide gene expression profiling. Clin Cancer Res 2005;11:2625 – 36.

3. von der Maase H, Sengelov L, Roberts JT, et al. Long-term survival results of a randomized trial comparing gemcitabine plus cisplatin, with methotrexate, vinblastine, doxorubicin, plus cisplatin in patients with bladder cancer. J Clin Oncol 2005;23:4602 – 8.

4. Vaughn DJ. Paclitaxel and carboplatin in bladder cancer: recent developments. Eur J Cancer 2000;36 Suppl 2:7 – 12.

5. Fechner G, Siener R, Reimann M, Kobalz L, Albers P. Randomised phase II trial of gemcitabine and paclitaxel second-line chemotherapy in patients with transitional cell carcinoma (AUO Trial AB 20/99). Int J Clin Pract 2006;60:27 – 31.

6. Titus B, Frierson HF, Jr., Conaway M, et al. Endothelin axis is a target of the lung metastasis suppressor gene RhoGDI2. Cancer Res 2005;65:7320 – 7.

7. Harding MA, Arden KC, Gildea JW, et al. Functional genomic comparison of lineage-related human bladder cancer cell lines with differing tumorigenic and metastatic potentials by spectral karyotyping, comparative genomic hybridization, and a novel method of positional expression profiling. Cancer Res 2002;62:6981 – 9.

8. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.

9. Ahmed SA, Gogal RM, Jr., Walsh JE. A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to [^{3}H]thymidine incorporation assay. J Immunol Methods 1994;170:211 – 24.

10. Nakayama GR, Caton MC, Nova MP, Parandoosh Z. Assessment of the Alamar Blue assay for cellular growth and viability *in vitro*. J Immunol Methods 1997;204:205 – 8.

11. Jain N, Thatte J, Braciale T, Ley K, O'Connell M, Lee JK. Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays. Bioinformatics 2003;19:1945 – 51.

12. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116 – 21.

13. Soukup M, Cho H, Lee JK. Robust classification modeling on microarray data using misclassification penalized posterior. Bioinformatics 2005;21 Suppl 1:i423 – 30.

14. Jebar AH, Hurst CD, Tomlinson DC, Johnston C, Taylor CF, Knowles

MA. FGFR3 and Ras gene mutations are mutually exclusive genetic events in urothelial cell carcinoma. Oncogene 2005;24:5218 – 25.

15. Sonpavde G, Petrylak DP. Perioperative chemotherapy for bladder cancer. Crit Rev Oncol Hematol 2006;57:133 – 44.

16. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531 – 7.

17. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics 2000;16: 906 – 14.

18. Nguyen DV, Rocke DM. Tumor classification by partial least squares using microarray gene expression data. Bioinformatics 2002;18:39 – 50.

19. Baksh S, Tommasi S, Fenton S, et al. The tumor suppressor RASSF1A and MAP-1 link death receptor signaling to Bax conformational change and cell death. Mol Cell 2005;18:637 – 50.

20. Matsuyama Z, Yanagisawa NK, Aoki Y, et al. Polyglutamine repeats of spinocerebellar ataxia 6 impair the cell-death-preventing effect of CaV2.1 Ca2 + channel-loss-of-function cellular model of SCA6. Neurobiol Dis 2004;17:198 – 204.

21. Bates S, Rowan S, Vousden KH. Characterisation of human cyclin G1 and G2: DNA damage inducible genes. Oncogene 1996;13:1103 – 9.

22. Corson TW, Huang A, Tsao MS, Gallie BL. KIF14 is a candidate oncogene in the 1q minimal region of genomic gain in multiple cancers. Oncogene 2005;24:4741 – 53.

23. Moufarij MA, Phillips DR, Cullinane C. Gemcitabine potentiates cisplatin cytotoxicity and inhibits repair of cisplatin-DNA damage in ovarian cancer cell lines. Mol Pharmacol 2003;63:862 – 9.

24. Pu YS, Chen J, Huang CY, Guan JY, Lu SH, Hour TC. Cross-resistance and combined cytotoxic effects of paclitaxel and cisplatin in bladder cancer cells. J Urol 2001;165:2082 – 5.

25. Zupi G, Scarsella M, D'Angelo C, et al. Potentiation of the antitumoral activity of gemcitabine and paclitaxel in combination on human breast cancer cells. Cancer Biol Ther 2005;4:866 – 71.