• Retracted •

nature medicine

Genomic signatures to guide the use of chemotherapeutics

Anil Potti[1,2], Holly K Dressman[1,3], Andrea Bild[1,3], Richard F Riedel[1,2], Gina Chan[4], Robyn Sayer[4], Janiel Cragun[4], Hope Cottrill[4], Michael J Kelley[2], Rebecca Petersen[5], David Harpole[5], Jeffrey Marks[5], Andrew Berchuck[1,6], Geoffrey S Ginsburg[1,2], Phillip Febbo[1–3], Johnathan Lancaster[4] & Joseph R Nevins[1–3]

Using *in vitro* drug sensitivity data coupled with Affymetrix microarray data, we developed gene expression signatures that predict sensitivity to individual chemotherapeutic drugs. Each signature was validated with response data from an independent set of cell line studies. We further show that many of these signatures can accurately predict clinical response in individuals treated with these drugs. Notably, signatures developed to predict response to individual agents, when combined, could also predict response to multidrug regimens. Finally, we integrated the chemotherapy response signatures with signatures of oncogenic pathway deregulation to identify new therapeutic strategies that make use of all available drugs. The development of gene expression profiles that can predict response to commonly used cytotoxic agents provides opportunities to better use these drugs, including using them in combination with existing targeted therapies.

Numerous advances have been achieved in the development, selection and application of chemotherapeutic agents, sometimes with remarkable clinical successes—as in the case of treatment for lymphomas or platinum-based therapy for testicular cancers[1]. In addition, in several instances, combination chemotherapy in the postoperative (adjuvant) setting has been curative. However, most people with advanced solid tumors will relapse and die of their disease. Moreover, administration of ineffective chemotherapy increases the probability of side effects, particularly those from cytotoxic agents, and of a consequent decrease in quality of life[1,2].

Recent work has demonstrated the value in using biomarkers to select individuals for various targeted therapeutics, including tamoxifen, trastuzumab and imatinib mesylate. In contrast, equivalent tools to select those most likely to respond to the commonly used chemotherapeutic drugs are lacking[3].

With the goal of developing genomic predictors of chemotherapy sensitivity that could direct the use of cytotoxic agents to those most likely to respond, we combined *in vitro* drug response data, together with microarray gene expression data, to develop models that could potentially predict responses to various cytotoxic chemotherapeutic drugs[4]. We now show that these signatures can predict clinical or pathologic response to the corresponding drugs, including combinations of drugs. We further use the ability to predict deregulated oncogenic signaling pathways in tumors to develop a strategy that identifies opportunities for combining chemotherapeutic drugs with targeted therapeutic drugs in a way that best matches the characteristics of the individual.

RESULTS

A gene expression–based predictor of sensitivity to docetaxel

To develop predictors of cytotoxic chemotherapeutic drug response, we used an approach similar to previous work analyzing the NCI-60 panel[4] from the US National Cancer Institute (NCI). We first identified cell lines that were most resistant or sensitive to docetaxel (**Fig. 1a**,**b**) and then genes whose expression correlated most highly with drug sensitivity, and used Bayesian binary regression analysis to develop a model that differentiates a pattern of docetaxel sensitivity from that of resistance. A gene expression signature consisting of 50 genes was identified that classified cell lines on the basis of docetaxel sensitivity (**Fig. 1b**, right).

In addition to leave-one-out cross-validation, we used an independent dataset derived from docetaxel sensitivity assays in a series of 30 lung and ovarian cancer cell lines for further validation. The significant correlation ($P < 0.01$, log-rank test) between the predicted probability of sensitivity to docetaxel (in both lung and ovarian cell lines) (**Fig. 1c**, left) and the respective 50% inhibitory concentration (IC_{50}) for docetaxel confirmed the capacity of the docetaxel predictor to predict sensitivity to the drug in cancer cell

[1]Duke Institute for Genome Sciences and Policy, Duke University, Box 3382, Durham, North Carolina 27710, USA. [2]Department of Medicine, Duke University Medical Center, Box 31295, Durham, North Carolina 27710, USA. [3]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Box 3054, Durham, North Carolina 27710, USA. [4]Division of Gynecologic Surgical Oncology, H. Lee Moffitt Cancer Center & Research Institute, University of South Florida, 12902 Magnolia Drive, Tampa, Florida 33612, USA. [5]Department of Surgery, Duke University Medical Center, Box 3627, Durham, North Carolina 27710, USA. [6]Department of Obstetrics and Gynecology, Duke University Medical Center, Box 3079, Durham, North Carolina 27710, USA. Correspondence should be addressed to J.R.N. (nevin001@mc.duke.edu).

Received 11 July; accepted 12 September; published online 22 October; corrected online 27 October 2006 (details online) and corrected after print 21 July 2008; doi:10.1038/nm1491

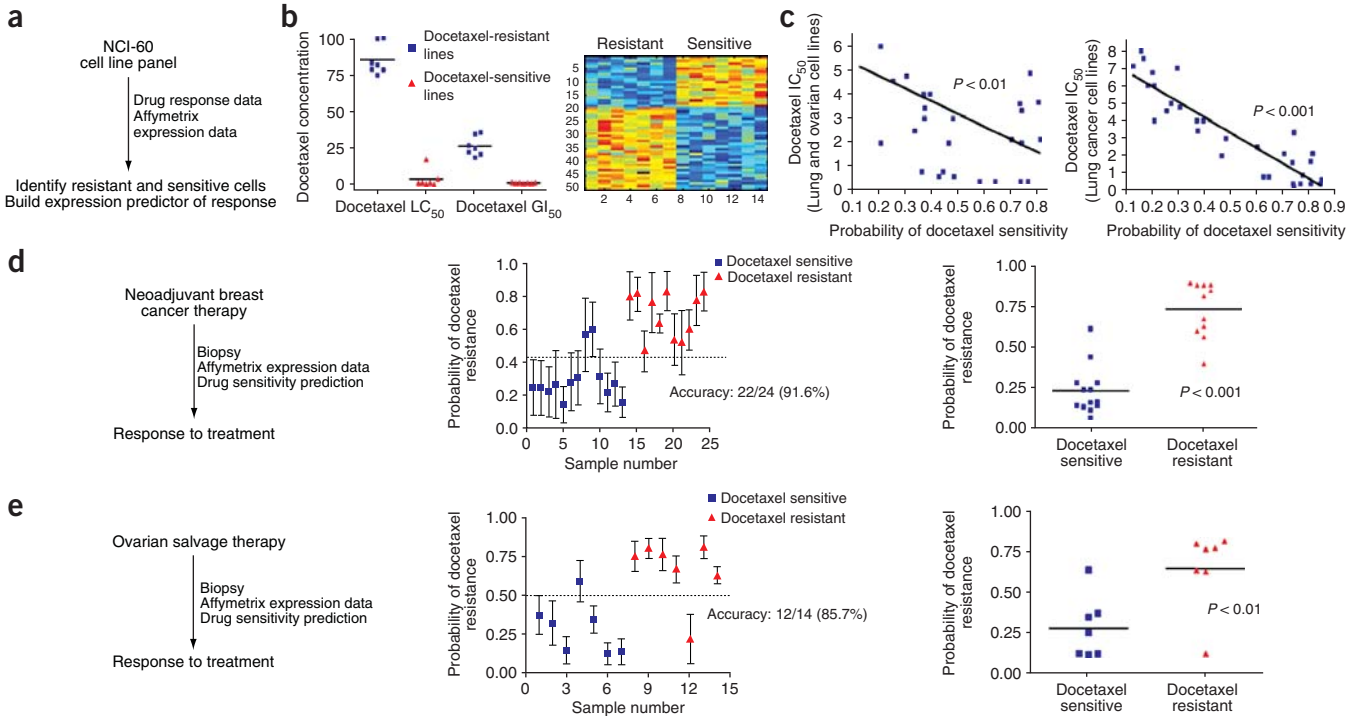

Figure 1 A gene expression signature that predicts sensitivity to docetaxel. (**a**) Strategy for generating the chemotherapeutic response predictor. (**b**) Left, cell lines from the NCI-60 panel used to develop the *in vitro* signature of docetaxel sensitivity. There is a statistically significant difference (Mann-Whitney *U*-test) in the IC_{50} and LC_{50} of the cell lines chosen to represent the sensitive and resistant subsets. Right, expression plots for genes selected for discriminating the docetaxel-resistant and docetaxel-sensitive NCI-60 cell lines, with blue representing the lowest expression and red the highest. Each column in the figure represents an individual sample. Each row represents an individual gene, ordered from top to bottom according to regression coefficient. (**c**) Left, validation of the docetaxel response prediction model in an independent set of lung and ovarian cancer cell line samples. A collection of lung and ovarian cell lines were used in a cell proliferation assay to determine the IC_{50} of docetaxel in the individual cell lines. Right, validation of the docetaxel response prediction model in another independent set of 29 lung cancer cell line samples. (**d**) Left, a strategy for assessment of the docetaxel response predictor as a function of clinical response in the breast neoadjuvant setting. Middle, predicted probability of docetaxel sensitivity in a collection of samples from a breast cancer single-agent neoadjuvant study. Right, a single-variable scatter plot of a significance test of the predicted probabilities of sensitivity to docetaxel in the sensitive and resistant tumors ($P < 0.001$, Mann-Whitney *U*-test). (**e**) Left, a strategy for assessment of the docetaxel response predictor as a function of clinical response in advanced ovarian cancer. Middle, predicted probability of docetaxel sensitivity in a collection of samples from a prospective single agent salvage therapy study. Right, a single-variable scatter plot showing statistical significance ($P < 0.01$, Mann-Whitney *U*-test).

lines (**Supplementary Fig. 1** online). For each set of cell lines, the accuracy exceeded 80%. Finally, we made use of a second independent dataset that measured docetaxel sensitivity in a series of 29 lung cancer cell lines (Gene Expression Omnibus (GEO) database (see URLs) accession number GSE4127). The docetaxel sensitivity model developed from the NCI-60 panel again predicted sensitivity in this independent dataset, also with an accuracy exceeding 80% ($P < 0.001$, log-rank test; **Fig. 1c**, right).

The expression signature predicts clinical docetaxel response

We used published studies with clinical and genomic data that linked gene expression data with clinical response to docetaxel in a breast cancer neoadjuvant study[5] (**Fig. 1d**) to test the capacity of the *in vitro* docetaxel sensitivity predictor to accurately identify those individuals that responded to docetaxel. Using a 0.45 predicted probability of response as the cutoff for predicting positive response, as determined by receiver operator characteristic curve analysis (**Supplementary Fig. 1**), the *in vitro*–generated profile correctly predicted docetaxel response in 22 of 24 clinical samples, for an overall accuracy of 91.6% (**Fig. 1d**). Applying a Mann-Whitney *U*-test for statistical significance demonstrated the capacity of the predictor to distinguish resistant from sensitive individuals (**Fig. 1d**, right). We extended this further by predicting the response to docetaxel as salvage therapy for individuals with ovarian cancer that was refractory to primary therapy (**Fig. 1e**), and the prediction achieved an accuracy exceeding 85% (**Fig. 1e**, middle). Further, an analysis of statistical significance demonstrated the capacity of the predictors to distinguish individuals with resistant versus sensitive disease (**Fig. 1e**, right).

We also performed a complementary analysis, using the clinical response data to generate a signature and then predicting sensitivity of NCI-60 cell lines to docetaxel (**Supplementary Fig. 1**). Genes represented in the initial *in vitro*–generated docetaxel predictor and the alternative *in vivo* predictor showed considerable overlap. Notably, both predictors were linked to expected targets for docetaxel, including *BCL2*, *WDR7* (also known as *TRAG*), *ERBB2* and tubulin genes, all previously described to be involved in taxane chemoresistance[6–9] (**Supplementary Table 1** online). We also noted that the predictor of docetaxel sensitivity developed from the NCI-60 data, using the approach described here, was more accurate in predicting response in the ovarian samples than a predictor developed from the breast neoadjuvant study data (85.7% versus 64.3%; **Supplementary Fig. 1**).

A panel of signatures that predict chemotherapeutic sensitivity

Given the development of a docetaxel response predictor, we examined the NCI-60 dataset for other opportunities to develop predictors

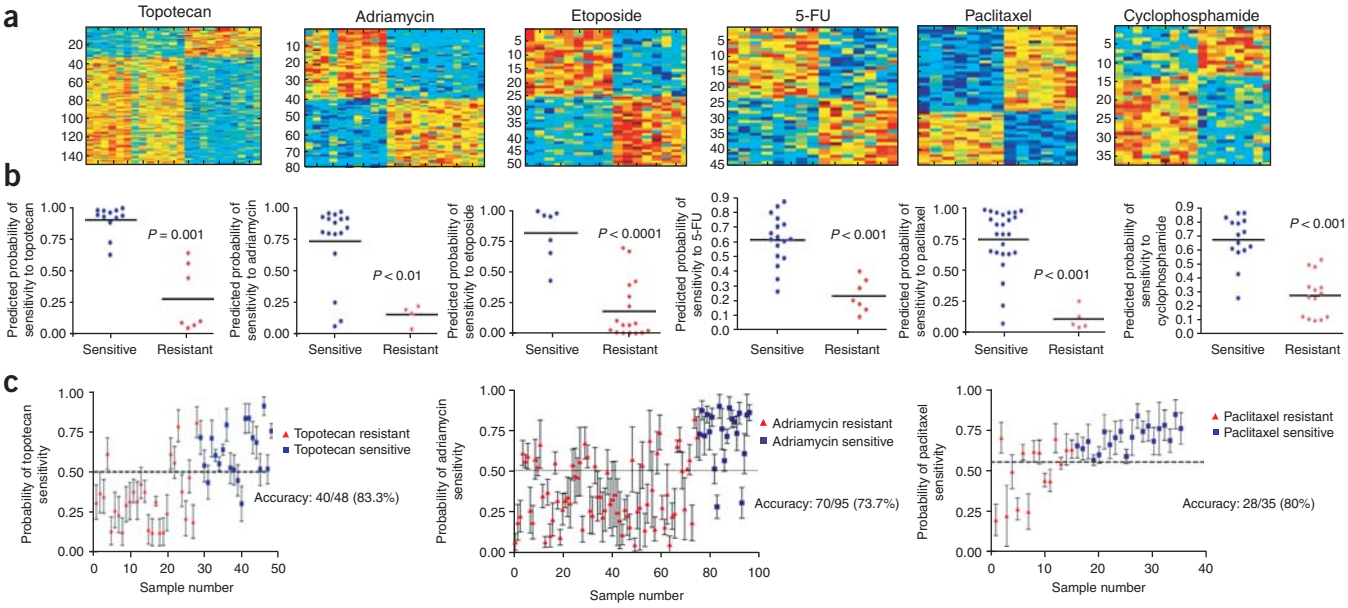

Figure 2 Development of a panel of gene expression signatures that predict sensitivity to chemotherapeutic drugs. (**a**) Gene expression patterns selected for predicting response to the indicated drugs. (**b**) Independent validation of the chemotherapy response predictors in an independent set of cancer cell lines that have dose-response and Affymetrix expression data[10]. A single-variable scatter plot of a significance test of the predicted probabilities of sensitivity to any given drug in the sensitive and resistant cell lines (*P* values, Mann-Whitney *U*-test). (**c**) Prediction of single-agent therapy response in clinical samples using *in vitro* cell line–based expression signatures of chemosensitivity. Left, the predicted probability of sensitivity to topotecan when compared to actual clinical response data ($n = 48$). Middle, the accuracy of the adriamycin predictor in a cohort of 122 samples. Right, predictive accuracy of the cell line–based paclitaxel predictor when the drug was used as a salvage chemotherapy in advanced ovarian cancer ($n = 35$).

of chemotherapeutic response. We developed a series of expression profiles from the NCI-60 dataset (**Fig. 2a**) that predict response to topotecan, adriamycin, etoposide, 5-fluorouracil (5-FU), paclitaxel and cyclophosphamide (Cytoxan). In each case, the leave-one-out cross-validation analyses demonstrated the capacity of these profiles to accurately predict the samples used in the development of the predictor (**Supplementary Fig. 2** online). Each profile was then further validated using *in vitro* response data from independent datasets; in each case, the profile developed from the NCI-60 data was capable of accurately (>85%) predicting response in the separate dataset of approximately 30 cancer cell lines for which the dose-response information and relevant Affymetrix U133A gene expression data were publicly available[10] (**Supplementary Fig. 2** and **Supplementary Table 2** online). Once again, applying a Mann-Whitney *U*-test for statistical significance demonstrated the capacity of the predictor to distinguish resistant from sensitive individuals (**Fig. 2b**).

We then evaluated the extent to which a signature was also specific for an individual chemotherapeutic agent. Using the validations of chemosensitivity seen in the independent European cell line data[10], it was clear that each signature is specific for the drug that was used to develop the predictor (**Supplementary Fig. 3** online).

Given the ability of the *in vitro*–developed gene expression profiles to predict response to docetaxel in the clinical samples, we extended this approach to test the ability of additional signatures to predict responses to commonly used salvage therapies for ovarian cancer and in another, independent dataset of samples cultured from adriamycin-treated individuals (GEO accession nos. GSE2351, GSE649, GSE650 and GSE651). Each of these predictors was capable of accurately predicting the response to the drugs in clinical samples, achieving an accuracy in excess of 74% overall (**Fig. 2c**). In each case, the positive and negative predictive values confirmed the validity of the approach (**Supplementary Table 2**).

Signatures predict response to multidrug regimens

Many therapeutic regimens make use of combinations of chemotherapeutic drugs. To evaluate whether individual signatures of response could predict response to regimens containing combinations of drugs, we analyzed data from a breast neoadjuvant treatment study that used a combination of paclitaxel, 5-FU, adriamycin and cyclophosphamide (TFAC)[11,12] (**Fig. 3a**). The available data from the 51 participants were used to predict response using each of the single-agent signatures (for paclitaxel, 5-FU, adriamycin and cyclophosphamide) developed from the NCI-60 cell line analysis. The predicted response based on each of the individual chemosensitivity signatures indicated a significant distinction between the responders ($n = 13$) and nonresponders ($n = 38$), with the exception of the prediction for 5-FU (**Fig. 3a**, middle). The sensitivity to the four agents in the TFAC preoperative (neoadjuvant) regimen was predicted as a combined probability. The prediction of response based on a combined probability of sensitivity built from the individual chemosensitivity predictions yielded a statistically significant ($P < 0.0001$, Mann-Whitney *U*-test) distinction between the responders and nonresponders (**Fig. 3a**, right).

As a further validation of the capacity to predict response to combination therapy, we analyzed gene expression data generated from a collection of breast cancer ($n = 45$) samples from subjects who received 5-FU, adriamycin and cyclophosphamide (FAC) as adjuvant chemotherapy (GEO database accession no. GSE3143). The predicted response based on the individual signatures indicated a significant distinction between the responders ($n = 34$) and nonresponders ($n = 11$; **Fig. 3b**, left). Furthermore, the combined probability of sensitivity to the three agents in the FAC regimen yielded a clear and significant ($P < 0.001$, Mann-Whitney *U*-test) distinction between the responders and nonresponders (accuracy 82.2%, positive predictive value 90.3%, negative predictive value 64.3%; **Fig. 3b**). Although it is difficult to interpret the prediction of clinical response in the adjuvant

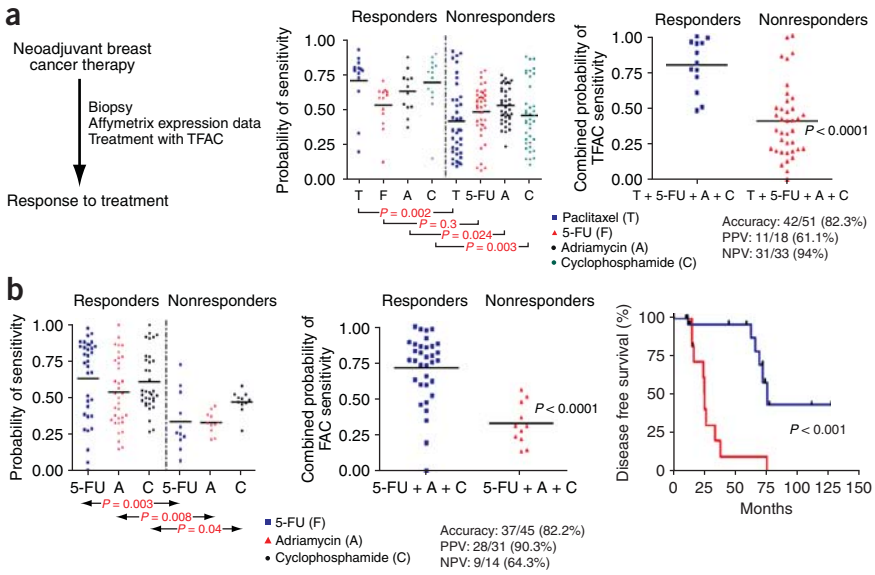

Figure 3 Prediction of response to combination therapy. (**a**) Left, strategy for assessment of chemotherapy response predictors in combination therapy as a function of pathologic response. Middle, prediction of clinical response to neoadjuvant chemotherapy involving paclitaxel, 5-FU, adriamycin and cyclophosphamide (TFAC) using the single-agent *in vitro* chemosensitivity signatures developed for each of these drugs. Right, prediction of response (38 nonresponders, 13 responders) using a combined probability predictor assessing the probabilities of all four chemosensitivity signatures in 51 people treated with TFAC chemotherapy ($P < 0.0001$, Mann-Whitney *U*-test). Response was defined as a complete pathologic response after completion of TFAC neoadjuvant therapy. (**b**) Left, prediction of clinical response ($n = 45$) to adjuvant chemotherapy involving 5-FU, adriamycin and cyclophosphamide (FAC) using the single-agent *in vitro* chemosensitivity predictors developed for these drugs. Middle, prediction of response (34 responders, 11 nonresponders) using a combined probability predictor assessing the probabilities of all four chemosensitivity signatures in 45 people treated with FAC chemotherapy. Right, Kaplan-Meier survival analysis for individuals predicted to be sensitive (blue curve) or resistant (red curve) to FAC adjuvant chemotherapy. PPV, positive predictive value; NPV, negative predictive value.

setting, as many of these people were probably free of disease after surgery, the accurate identification of nonresponders is a clear endpoint that does confirm the capacity of the signatures to predict clinical response. We also examined the prognostic significance of the prediction of response to FAC: it defined a poor prognosis group (**Fig. 3b**, right). Taking these results together, we conclude that the signatures of chemosensitivity generated from the NCI-60 panel do indeed have the capacity to predict therapeutic response in individuals receiving either single agent or combination chemotherapy (**Supplementary Tables 2** and **3** online).

An examination of the genes that constituted the paclitaxel predictor identified microtubule-associated protein tau (*MAPT*), described previously as a determinant of paclitaxel sensitivity[12]. Also consistent with previous reports[5–7], *TP53*, methylenetetrahydrofolate reductase (*MTHFR*) and DNA repair genes constituted the 5-FU predictor, and excision repair mechanism genes (for example, *ERCC4*), retinoblastoma pathway genes and *BCL2* constituted the adriamycin predictor (**Supplementary Table 1**).

Patterns of predicted chemotherapy response across tumor types

The panel of chemotherapy response predictors described in **Figure 3** was used to profile the potential options for use of these drugs, by predicting the likelihood of sensitivity to the seven agents in a large collection of breast, lung and ovarian tumor samples. We then clustered the samples according to patterns of predicted sensitivity to the various chemotherapeutics and plotted a heatmap. There were clearly evident patterns of predicted sensitivity to the various agents (**Fig. 4**). In many cases, the predicted sensitivities to the chemotherapeutic agents were consistent with the previously documented efficacy of single-agent chemotherapies in the individual tumor types[13]. For instance, the predicted response rates for etoposide, adriamycin, cyclophosphamide and 5-FU approximated the observed response for these single agents in individuals with breast cancer (**Supplementary Fig. 3**). Likewise, the predicted sensitivities to etoposide, docetaxel and paclitaxel approximated the observed response for these single agents in individuals with lung cancer (**Supplementary Fig. 3**). This analysis also suggested possibilities for alternate treatments. As an example, it would seem that individuals with breast cancer who are likely to respond to 5-FU are resistant to adriamycin and docetaxel (**Supplementary Fig. 4** online). Likewise, in lung cancer, docetaxel-sensitive individuals are likely to be resistant to etoposide (**Supplementary Fig. 4**). This is a potentially useful observation, considering that both etoposide and docetaxel are viable front-line options (in conjunction with cisplatin or carboplatin) for people with lung cancer[1]. A similar relationship is seen between topotecan and adriamycin, both of which are agents used in salvage chemotherapy for ovarian cancer (**Supplementary Fig. 4**). Thus, by identifying people or cohorts resistant to a certain standard-of-care agent (for example, topotecan), one could, by choosing an alternative standard of care agent (for example, adriamycin), avoid the side effects of the former agent without compromising outcome.

Linking chemotherapy sensitivity to oncogenic pathway status

Most people who are resistant to chemotherapeutic agents are recruited into a second- or third-line therapy or enrolled in a clinical trial[14,15]. Moreover, even those who initially respond to a given agent are likely to eventually suffer a relapse, and thus, in either case, additional therapeutic options are needed. As one approach to identifying such options, we have taken advantage of our recent work that describes the development of gene expression signatures that reflect the activation of several oncogenic pathways[16] (**Supplementary Fig. 5** and **Supplementary Table 1** online). To illustrate this approach, we first stratified the NCI cell lines based on predicted docetaxel response and then examined the patterns of pathway deregulation associated with docetaxel sensitivity or resistance (**Supplementary Fig. 5**). Regression analysis showed a significant relationship between phosphatidylinositol 3-OH (PI3)-kinase pathway deregulation and docetaxel resistance (**Supplementary Fig. 5**).

These results indicated an opportunity to use a PI3-kinase inhibitor in this subgroup, given our recent observations that have demonstrated a linear positive correlation between the probability of pathway deregulation and sensitivity to targeted drugs[16]. To address this directly, we predicted docetaxel sensitivity and probability of oncogenic pathway deregulation using DNA microarray data from 17 non–small cell lung cancer cell lines (**Fig. 5a**, left). Consistent with

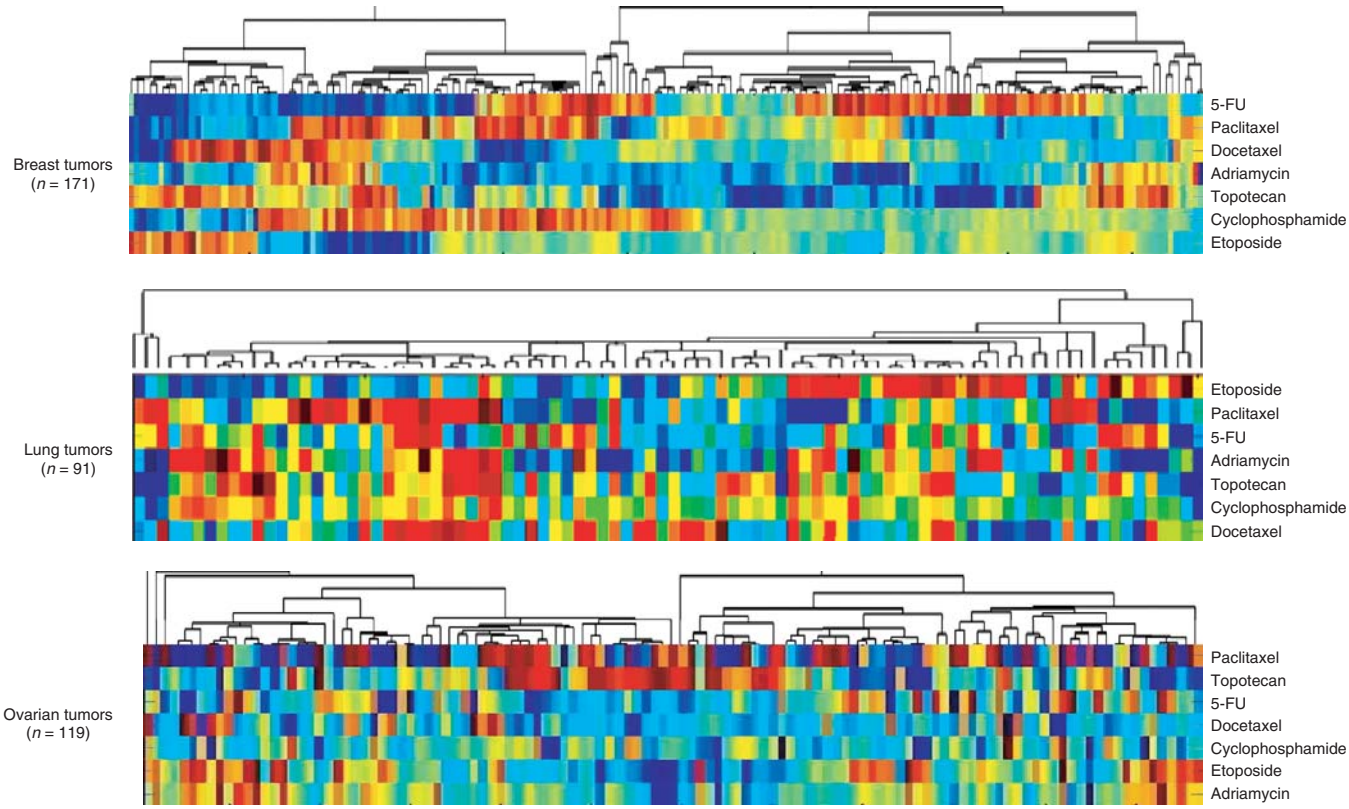

Figure 4 Patterns of predicted sensitivity to common chemotherapeutic drugs in human cancers. Hierarchical clustering of a collection of breast ($n = 171$), lung ($n = 91$) and ovarian cancer ($n = 119$) samples according to patterns of predicted sensitivity to the various chemotherapeutics. Predictions were plotted as a heatmap in which high probability of sensitivity, or response, is indicated by red and low probability, or resistance, is indicated by blue.

the analysis of the NCI-60 cell line panel, the cell lines predicted to be resistant to docetaxel were also predicted to show PI3-kinase pathway activation ($P = 0.03$, log-rank test; **Supplementary Fig. 5**). In parallel, the lung cancer cell lines were assayed for sensitivity to a PI3-kinase–specific inhibitor (LY-294002), using a standard measure of cell proliferation[14–16]. The cell lines showing an increased probability of PI3-kinase pathway activation were also more likely to respond to the PI3-kinase inhibitor ($P = 0.001$, log-rank test; **Fig. 5b**, left). The same relationship held for prediction of resistance to docetaxel: these cells were more likely to be sensitive to PI3-kinase inhibition ($P < 0.001$, log-rank test; **Fig. 5b**, left).

An analysis of a panel of ovarian cancer cell lines provided a second example of an opportunity to link predictions of chemotherapy sensitivity to oncogenic pathway deregulation. Ovarian cell lines that were predicted to be topotecan resistant (**Fig. 5a**, right) had a higher likelihood of Src pathway deregulation, and there was a significant linear relationship ($P = 0.001$, log-rank test) between the probability of topotecan resistance and sensitivity to a drug (SU6656) that inhibits the Src pathway (**Fig. 5b**, right). The results of these assays clearly demonstrated an opportunity to mitigate drug resistance (for example, resistance to docetaxel or topotecan) using a specific pathway–targeted agent (PI3-kinase or Src inhibitor, respectively).

Taken together, these data demonstrate a rational approach to the identification of therapeutic options for chemotherapy-resistant individuals, as well as to the identification of new combinations for chemotherapy-sensitive individuals, and thus have the potential to change the current paradigm of cancer care. Prospective validation studies are, however, needed to confirm the effectiveness of this approach.

DISCUSSION

The practice of oncology continually faces the challenge of matching the right therapeutic regimen with the right individual, balancing relative benefit with risk to achieve the most favorable outcome. This challenge is often daunting, with marginal success rates in many advanced disease contexts—probably reflecting the enormous complexity of the disease process coupled with an inability to properly guide the use of available therapeutics[1,2,15]. The results we present here, focusing on the development of signatures that predict sensitivity to common cytotoxic chemotherapeutic drugs, address this limitation and have the potential to identify drugs—and combinations of drugs—that best match the individual. The experimental strategy for analyses used in this study is similar to that used for the development of oncogenic pathway signatures[16]: samples representing extreme cases are used to train the expression data to develop a signature that can predict drug sensitivity. Of note, we further demonstrated that these signatures can predict clinical drug response.

The importance of selecting individuals likely to respond to a given therapeutic agent is perhaps best illustrated by the example of trastuzumab. In the absence of selection, the overall response rate in people with breast cancer is approximately 10%. In contrast, for those selected on the basis of Her2 amplification, the overall response rate rises to 35–50%[17]. We suggest that the gene expression signatures predicting response to various cytotoxic chemotherapeutic agents may provide an opportunity to optimize the use of these drugs.

Previous work has described the use of gene expression data, coupled with *in vitro* drug sensitivity assays, to develop signatures that could be used to classify response to therapy[18,19]. This involved

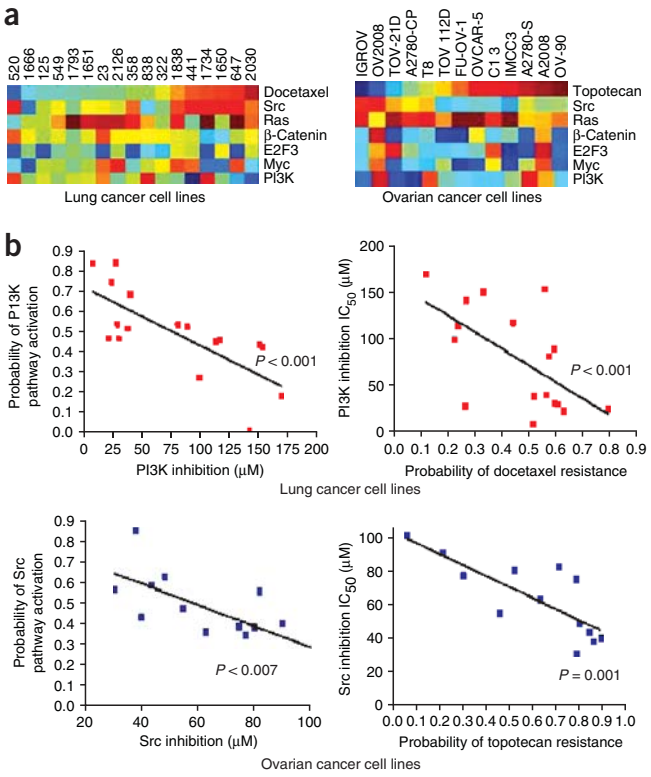

Figure 5 Relationship between predicted chemotherapeutic sensitivity and oncogenic pathway deregulation. (**a**) Left, probability of oncogenic pathway deregulation as a function of predicted docetaxel sensitivity in a series of lung cancer cell lines. Right, probability of oncogenic pathway deregulation as a function of predicted topotecan sensitivity in a series of ovarian cancer cell lines (red, sensitive; blue, resistant). (**b**) Top, lung cancer cell lines showing an increased probability of PI3-kinase (PI3K) activation were also more likely to respond to a PI3-kinase inhibitor ($P = 0.001$, log-rank test), as measured by sensitivity to the drug in assays of cell proliferation. Furthermore, those cell lines predicted to be resistant to docetaxel were more likely to be sensitive to PI3-kinase inhibition ($P < 0.001$, log-rank test). Bottom, the relationship between Src pathway deregulation and topotecan resistance in a set of 13 ovarian cancer cell lines. Ovarian cell lines that are predicted to be topotecan resistant (**a**, right) have a higher likelihood of Src pathway deregulation and there is a significant linear relationship ($P = 0.001$, log-rank test) between the probability of topotecan resistance and sensitivity to a drug that inhibits the Src pathway.

the development of signatures that could identify chemoresistance to the combination of drugs used in the treatment of acute lymphoblastic leukemia. Our findings that individual signatures can accurately predict response to drugs in both neoadjuvant and adjuvant chemotherapy settings, where combinations of cytotoxic drugs are used, clearly demonstrate the utility of the approach. We also describe a rational approach to identifying new drug regimens, making use of previously described signatures of oncogenic pathway deregulation[16], to provide a potential strategy for the development of therapeutic regimens that would be based on knowledge of the status of an individual's tumor.

Finally, we note that the availability of predictors of chemotherapy response, which have been shown to predict clinical response, provides an opportunity to apply these predictors in present day practice. There are several instances in which people are treated with one or more therapeutic regimens that each have equal efficacy. An ability to select

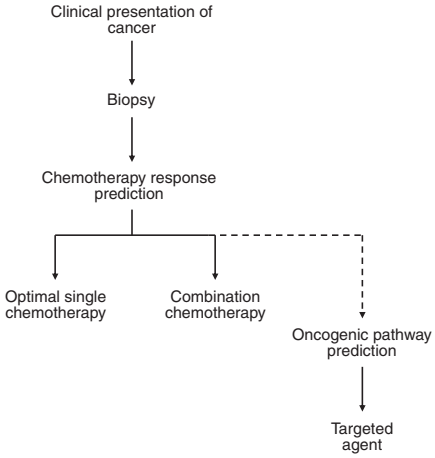

Figure 6 Scheme for using chemotherapeutic and oncogenic pathway predictors to identify individualized therapeutic options.

the most effective therapy from among a panel of standard-of-care agents is a strategy that can be easily applied[20]. Indeed, we believe this represents an opportunity for a clinical trial that would evaluate the performance of 'random' selection of agents (current practice) versus a genomic signature–based selection as an initial step in the process of achieving a more individualized treatment strategy (**Fig. 6**). Notably, such a study may also identify those individuals likely to be resistant to the available standard-of-care drugs, opening the way to a second-generation trial that could evaluate the most effective treatment for these people. Ultimately, we suggest that future treatment strategies might be based on an analysis of an individual's tumor, which would then allow the development of a profile of likely sensitivity to common chemotherapeutic drugs as well as to targeted therapies. Based on this information, each individual might then be assigned to a combination regimen that best matches the profile from the tumor.

METHODS

NCI-60 data. The complete details of the methods involved in the development of the individual chemosensitivity predictors are available in **Supplementary Methods** online. Briefly, the $-\log_{10}$ values of the 50% growth inhibition doses (IC_{50}; also known as GI_{50}), total growth inhibition doses and 50% cytotoxic doses (LC_{50}) data were used to populate a matrix in MATLAB software, along with the relevant expression data for the individual cell lines. Where multiple entries for a drug screen existed (by NCS number), the entry with the largest number of replicates was included. Incomplete data were assigned as 'not a number' for statistical purposes. To develop an *in vitro* gene expression–based predictor of sensitivity and resistance from the pharmacologic data used in the NCI-60 drug screen studies, we chose cell lines within the NCI-60 panel that would represent the extremes of sensitivity to a given chemotherapeutic agent (mean $IC_{50} \pm 1$ s.d.). Relevant expression data (revised data gathered using the Affymetrix U95A2 GeneChip) for the solid tumor cell lines and the respective pharmacological data for the chemotherapeutics were downloaded from the NCI website. The individual drug sensitivity and resistance data from the selected solid tumor NCI-60 cell lines were then used in a supervised analysis using binary regression methodologies, as described previously[21], to develop models predictive of chemotherapeutic response.

Human ovarian cancer samples. We measured expression of 22,283 genes in 13 ovarian cancer cell lines and 119 advanced (Federation Internationale de Gynecologie et d'Obstetrique (FIGO) stage III or IV) serous epithelial ovarian carcinomas using Affymetrix U133A GeneChips. All ovarian cancers were obtained at initial cytoreductive surgery. All tissues were collected under the

auspices of respective institutional (Duke University Medical Center and H. Lee Moffitt Cancer Center) review board–approved protocols involving written informed consent.

Full details of the methods used for RNA extraction and development of gene expression signatures representing deregulation of oncogenic pathways in the tumor samples were recently described[16] and are available in **Supplementary Methods**. Response to therapy was evaluated using standard criteria for those with measurable disease, based upon World Health Organization guidelines[22]. Complete details are available in **Supplementary Methods**.

Validation of *in vitro* chemotherapy response signatures. Complete details of the publicly available datasets used for independent validation of the genomic signatures of chemosensitivity are described in **Supplementary Methods**.

PI3-kinase signature. A signature representative of PI3-kinase activation was developed as previously described[16]. The genes that constitute the PI3-kinase signature are shown in **Supplementary Table 1**.

Lung and ovarian cancer cell culture. Total RNA was extracted and oncogenic pathway predictions were performed similarly to the methods described previously[16]. Complete details are available in **Supplementary Methods**.

Cross-platform Affymetrix Gene Chip comparison. To map the probe sets across various generations of Affymetrix GeneChip arrays, we used an in-house program, Chip Comparer (see URLs), as described previously[16].

Cell proliferation assays. Methods for drug sensitivity assays are described in **Supplementary Methods**.

Statistical analysis methods. Analysis of expression data was performed as previously described[16,21,23,24] and is detailed in **Supplementary Methods**. In instances where a combined probability of sensitivity to a combination chemotherapeutic regimen was required based on the individual drug sensitivity patterns, we used the probabilities of response to individual drugs and used the theorem for combined probabilities as described by William Feller to deduce a probability of response to a combination of the drugs being studied. The result was then mean-centered to give a probability between 0 and 1. Hierarchical clustering of tumor predictions was performed using Gene Cluster 3.0 (ref. 25). Genes and tumors were clustered using average linkage with the uncentered correlation similarity metric. Standard linear regression analyses and their significance (log-rank test) were generated for the drug response data and for correlation between drug response and probability of chemosensitivity or pathway deregulation using GraphPad software.

URLs. NCI, http://dtp.nci.nih.gov/docs/cancer/cancer_data.html. GEO database, http://www.ncbi.nlm.nih.gov/geo/. Chip Comparer, http://tenero.duhs.duke.edu/genearray/perl/chip/chipcomparer.pl.

Note: Supplementary information is available on the Nature Medicine website.

ACKNOWLEDGMENTS

The authors would like to acknowledge and thank L. Pusztai for providing us with the clinical and expression data on the samples studied from the M.D. Anderson Cancer Center. Funding for this work was supplied in part by research grants from the US National Institutes of Health (NCI-U54 CA112952-02 (J.R.N.) and R01CA106520 (J.R.N.)). A.P. is an American Association for Cancer Research Translational Research Grant awardee, J.L. receives research support from Glaxo-Smith-Kline and Sanofi-Aventis Pharmaceuticals and Orthobiotech, and P.F. is a Damon Runyon Cancer Research Investigator. Additional sources of funding included the Ovarian Cancer Research Fund and the Hearing the Ovarian Cancer Whisper Jacquie Liggett Fellowship.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Medicine* website for details).

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

1. Herbst, R.S. *et al.* Clinical cancer advances 2005: major research advances in cancer treatment, prevention, and screening—a report from the American Society of Clinical Oncology. *J. Clin. Oncol.* **24**, 190–205 (2006).
2. Breathnach, O.S. *et al.* Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: sobering results. *J. Clin. Oncol.* **19**, 1734–1742 (2001).
3. Blagosklonny, M.V. Why therapeutic response may not prolong the life of a cancer patient: selection for oncogenic resistance. *Cell Cycle* **4**, 1693–1698 (2005).
4. Staunton, J.E. *et al.* Chemosensitivity prediction by transcriptional profiling. *Proc. Natl. Acad. Sci. USA* **98**, 10787–10792 (2001).
5. Chang, J.C. *et al.* Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* **362**, 362–369 (2003).
6. Emi, M., Kim, R., Tanabe, K., Uchida, Y. & Toge, T. Targeted therapy against Bcl-2-related proteins in breast cancer cells. *Breast Cancer Res.* **7**, R940–R952 (2005).
7. Takahashi, T. *et al.* Cyclin A-associated kinase activity is needed for paclitaxel sensitivity. *Mol. Cancer Ther.* **4**, 1039–1046 (2005).
8. Modi, S. *et al.* Phosphorylated/activated HER2 as a marker of clinical resistance to single agent taxane chemotherapy for metastatic breast cancer. *Cancer Invest.* **23**, 483–487 (2005).
9. Langer, R. *et al.* Association of pretherapeutic expression of chemotherapy-related genes with response to neoadjuvant chemotherapy in Barrett carcinoma. *Clin. Cancer Res.* **11**, 7462–7469 (2005).
10. Gyorffy, B. *et al.* Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. *Int. J. Cancer* **118**, 1699–1712 (2005).
11. Rouzier, R. *et al.* Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin. Cancer Res.* **11**, 5678–5685 (2005).
12. Rouzier, R. *et al.* Microtubule-associated protein tau: a marker of paclitaxel sensitivity on breast cancer. *Proc. Natl. Acad. Sci. USA* **102**, 8315–8320 (2005).
13. DeVita, V.T., Hellman, S. & Rosenberg, S.A. *Cancer: Principles and Practice of Oncology* (Lippincott-Raven, Philadelphia, 2005).
14. Broxterman, H.J. & Georgopapadakou, N.H. Anticancer therapeutics: "addictive" targets, multi-targeted drugs, new drug combinations. *Drug Resist. Updat.* **8**, 183–197 (2005).
15. Minna, J.D., Gazdar, A.F., Sprang, S.R. & Herz, J. Cancer: a bull's eye for targeted lung cancer therapy. *Science* **304**, 1458–1461 (2004).
16. Bild, A. *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* **439**, 353–357 (2006).
17. Vogel, C.L. *et al.* Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J. Clin. Oncol.* **20**, 719–726 (2002).
18. Lugthart, S. *et al.* Identification of genes associated with chemotherapy crossresistance and treatment response in childhood acute lymphoblastic leukemia. *Cancer Cell* **7**, 375–386 (2005).
19. Holleman, A. *et al.* Gene expression patterns in drug resistant acute lymphoblastic leukemia cells and response to treatment. *N. Engl. J. Med.* **351**, 533–542 (2004).
20. Sargent, D.J., Conley, B.A., Allegra, C. & Collette, L. Clinical trial designs for predictive marker validation in cancer treatment trials. *J. Clin. Oncol.* **23**, 2020–2027 (2005).
21. Pittman, J., Huang, E., Wang, Q., Nevins, J.R. & West, M. Bayesian analysis of binary prediction tree models for retrospectively sampled outcomes. *Biostatistics* **5**, 587–601 (2004).
22. Miller, A.B., Hoogstraten, B., Staquet, M. & Winkler, A. Reporting results of cancer treatment. *Cancer* **47**, 207–214 (1981).
23. West, M. *et al.* Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc. Natl. Acad. Sci. USA* **98**, 11462–11467 (2001).
24. Ihaka, R. & Gentleman, R. A language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299–314 (1996).
25. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868 (1998).

Retraction: Genomic signatures to guide the use of chemotherapeutics

Anil Potti, Holly K Dressman, Andrea Bild, Richard F Riedel, Gina Chan, Robyn Sayer, Janiel Cragun, Hope Cottrill, Michael J Kelley, Rebecca Petersen, David Harpole, Jeffrey Marks, Andrew Berchuck, Geoffrey S Ginsburg, Phillip Febbo, Johnathan Lancaster & Joseph R Nevins

***Nat. Med.* 12, 1294–1300 (2006); published online 22 October 2006; corrected online 27 October 2006, 10 May 2007 and 10 October 2007 and corrected after print 21 July 2008; retracted 7 January 2011**

We wish to retract this article because we have been unable to reproduce certain crucial experiments showing validation of signatures for predicting response to chemotherapies, including docetaxel and topotecan. Although we believe that the underlying approach to developing predictive signatures is valid, a corruption of several validation data sets precludes conclusions regarding these signatures. As these results are fundamental to the conclusions of the paper, we formally retract the paper. We deeply regret the impact of this action on the work of other investigators.

Nature Medicine would also like to note that several of the earlier correction dates were either omitted or incorrect. The corrigenda published online 10 May 2007, 10 October 2007 and 21 July 2008 mistakenly omitted the earlier correction date of 27 October 2006. The correction in July 2008 went online on 21 July 2008 but was incorrectly noted in the corrigendum as having gone online 18 July 2008.