

Research Article

Development of Candidate Genomic Markers to Select Breast Cancer Patients for Dasatinib Therapy

Stacy Moulder^{1,2}, Kai Yan³, Fei Huang⁵, Kenneth R. Hess³, Cornelia Liedtke⁶, Feng Lin⁴, Christos Hatzis⁷, Gabriel N. Hortobagyi¹, W. Fraser Symmans⁴, and Lajos Pusztai¹

Abstract

Patient selection is important for targeted therapies, yet phase I/II trials are often underpowered for developing predictors of drug response. The goal of this research was to define genomic predictors for dasatinib that could be prospectively tested in early-phase clinical trials.

Gene expression profiles of dasatinib-sensitive and dasatinib-resistant cell lines ($n = 23$) were compared to develop a dasatinib-sensitivity index (modified DS index). A Src pathway activity index (revised Src index) was defined using genes induced by the Src transfection of mammary epithelial cells and was optimized to be reproducible across cell lines and human specimens. A dasatinib target index was devised using the weighted sum of 19 kinases that bind to dasatinib with variable affinity. The performance of these prediction models was assessed in independent cell lines with known dasatinib sensitivity. The feasibility of applying these genomic tests to human samples was evaluated on 133 biopsies of primary breast cancers.

The modified DS index showed 90% accuracy in independent breast cancer cell lines ($n = 12$) and the target index, but not the revised Src index signature, also distinguished dasatinib-sensitive and dasatinib-resistant cells ($P = 0.0024$). The genomic predictors showed acceptable reproducibility in replicate cell line and human gene expression data. When all three predictors were applied to the same 133 patient samples, the predictors identified different patient subsets as potentially sensitive.

We defined three conceptually different potential predictors of dasatinib response that were reproducible across cell lines and human data. These candidate markers are being tested in a clinical trial to determine their utility. *Mol Cancer Ther*; 9(5); 1120–7. ©2010 AACR.

Introduction

Implementation of personalized medicine requires transition from a “one drug treats all” philosophy into a more tailored approach to drug development involving enrichment of treatment cohorts with patients most likely to have therapeutic benefit. In breast cancer, gene expression profiles are emerging as tumor classifiers, prognostic indicators, and potential predictors of response to chemotherapy and hormonal therapies (1–8). These early successes motivate attempts to use this technology to develop treatment response predictors for novel

therapies in the context of phase I/II trials. Unfortunately, the small sample size and usually low tumor response rates limit the power of discovery of genomic predictors using tissue samples from patients treated in phase I/II trials (9). Recent results also suggest that the commonly used empirical discovery approach that relies on comparing molecular profiles of known responders with nonresponders may not yield reliable predictors for a variety of reasons. The multiple comparison problem inherent to microarray analysis is compounded by the small sample size of a typical phase II study and can lead to high rates of false discovery. The coordinated expression of thousands of genes associated with phenotypic characteristics of breast and other cancers can lead to the discovery of clinical phenotype predictor rather than a drug-specific response predictor (9).

In light of these limitations, it is reasonable to test if a more mechanistic approach to the development of response predictors may result in a higher rate of success. We previously proposed a clinical trial design to prospectively assess candidate response predictors using a tandem phase II two-stage evaluation design (9). Response predictors for clinical testing can be rationally developed based on data from preclinical cell line models or the known mechanism of action of the drug. These predictors have to be fully defined, including decision thresholds, and

Authors' Affiliations: Departments of ¹Breast Medical Oncology, ²Investigational Cancer Therapeutics, ³Bioinformatics/Quantitative Sciences, and ⁴Pathology of the University of Texas M.D. Anderson Cancer Center, Houston, Texas; ⁵Bristol-Myers Squibb Co., Princeton, New Jersey; ⁶University of Muenster, Muenster, Germany; and ⁷Nuvera Biosciences, Inc., Woburn, Massachusetts

Note: Supplementary material for this article is available at Molecular Cancer Therapeutics Online (<http://mct.aacrjournals.org/>).

S. Moulder and K. Yan are co-first authors.

Corresponding Authors: Stacy Moulder or Lajos Pusztai, Breast Medical Oncology, Unit 1354, The University of Texas M.D. Anderson Cancer Center, P.O. Box 301438, Houston, TX 77030. Phone: 713-792-2817; Fax: 713-794-4385. E-mail: smoulder@mdanderson.org or lpusztai@mdanderson.org

doi: 10.1158/1535-7163.MCT-09-1117

©2010 American Association for Cancer Research.

technically robust before they can be proposed for clinical evaluation to assess their predictive value in patients.

Dasatinib is an inhibitor of the bcr/abl kinase as well as several other protein kinases and is approved by the U.S. Food and Drug Administration for the treatment of chronic myeloid leukemia that relapse after imatinib and for the treatment of Philadelphia chromosome-positive acute lymphoblastic leukemia. This drug is also in clinical testing for the treatment of metastatic breast cancer. In a preliminary report of a recent study that included 43 patients with triple receptor-negative breast cancer who received single-agent dasatinib partial response was seen in only 2 of 36 patients (10). Two additional patients had stable disease for 4 months, for an overall clinical benefit rate of 9.3%. This study suggests that dasatinib is a potentially active drug for a small subset of patients but because of the overall low efficacy, patient selection to enrich for responders is critically important for future development. In the current study, we describe the development of three conceptually distinct genomic response predictors for dasatinib and report the reproducibility and feasibility of applying these tests to human gene expression data. These predictors are now tested for their clinical utility in a prospective multiarm clinical trial in metastatic breast cancer (11).

Dasatinib inhibits Src and several other members of this kinase family that have been implicated in breast cancer biology. Transfection of Src into primary mammary epithelial cells (HMEC) leads to the induction of several genes that can be considered as a Src activity signature (12). One could hypothesize that the greater the expression of these genes, the greater the Src activity and these tumors may be particularly sensitive to dasatinib. Investigators have also compared the gene expression profiles of dasatinib-sensitive to dasatinib-resistant breast cancer cell lines and generated a cell line-based genomic predictor of dasatinib response (13). Dasatinib also binds to many protein kinases other than Src, and these kinases may also contribute important biological functions to sustain malignant growth. We hypothesized that the higher the expression of these targets, the greater the sensitivity to dasatinib may be. Based on this reasoning, we propose a 19-gene dasatinib target index (T index) that represents the weighted sum of these genes weighed by the binding affinity of dasatinib to each kinase (14). In this article, we describe the transformation of these concepts into well-defined genomic predictors that can be applied to the human gene expression data. We also assess the reproducibility of prediction results in replicate clinical specimens and define the necessary thresholds for patient selection in a clinical trial.

Materials and Methods

Human gene expression data

All gene expression data were generated with Affymetrix HG-U133A chips as previously described (8). The raw gene expression profiles of 133 human breast cancers obtained at the time of diagnosis from fine needle biopsy

specimens of the cancer are available at the Web site of Department of Bioinformatics, M.D. Anderson Cancer Center (15). To measure reproducibility, 39 patient samples obtained by fine needle aspiration (FNA) were split and profiled twice, either in different laboratories or several months apart in the same laboratory (technical replicates). Finally, to determine the effect of sample acquisition methods (FNA versus core needle biopsy), data generated from genetic profiling of 37 matched pairs of tissue samples obtained from primary breast carcinomas using both FNA and core biopsy were analyzed.

Cell line gene expression data

The Src index was developed using gene expression data from HMECs transfected to express Src *in vitro*. (12) Src-transformed cells were compared using the Affymetrix HG-U133Plus2 array to control cells transfected with green fluorescent protein to identify the genes that induced the overexpression of Src. We downloaded the original data set and identified the 10 HMECs overexpressing Src and those who expressed green fluorescent protein (control) from the National Center for Biotechnology Information Gene Expression Omnibus Web site accession number GSE3141 to identify the Src-induced genes.

The previously published DS index was developed by categorizing 23 breast cancer cell lines using the mean \log_{10} IC₅₀ values (0.6 μ mol/L of dasatinib) to assign either a dasatinib-sensitive (IC₅₀ below the mean) or dasatinib-resistant (IC₅₀ above the mean) category to each cell line (13). Affymetrix U133 gene expression data from these cells are available at the Gene Expression Omnibus database under accession number GSE6569. Normalized data were used to select informative probe sets using unequal variance *t* test. This analysis identified 142 genes that were used to create the DS index. A separate 12 breast cancer and 23 lung cancer cell lines with known dasatinib sensitivity (IC₅₀ below the mean, sensitive) were used as an independent *in vitro* validation set and were provided by Bristol-Myers Squibb Company.

We also examined the robustness and stability of cell line-based gene signatures in cells grown in different laboratories. Access to two other cell line gene expression data sets, which were from cells not treated with dasatinib, was available for this purpose, including one set of data from the M.D. Anderson Cancer Center Breast Cancer Pharmacogenomic Laboratory ($n = 19$ cell lines; ref. 15) and another from a previous publication ($n = 51$ cell lines; refs. 16, 17). The two independent cell line data sets and the training set of 23 cell lines for the DS index had 14 cell lines in common. These 14 cell lines were used to assess the stability of prediction results across data sets generated in different laboratories from the same cell line.

Data preprocessing

All cell line and human gene expression data were normalized with the Microarray Suite version 5.0 using the global scaling. The trimmed mean target intensity of each array was arbitrarily set to 1,500. Normalized

gene expression values were transformed to the log₂ scale for further analysis.

Definition of informative probe sets

Two predictors, the Src Index and the cell line-derived sensitivity predictor, were developed using previously published data (12, 13).

Modified Src index. The previously published Src index was developed using HMECs transfected to express a variety of oncogenes, including Src (12). Gene expression profiles of Src-transformed cells were compared with parental cells transfected with green fluorescent protein to define a Src pathway signature. The original signature consisted of 73 probe sets from the Affymetrix HG-U133Plus2 array; however, only 46 of the 73 probe sets are present in the Affymetrix HG-133A arrays used in our human samples; therefore, the gene signature for our modified Src pathway index is based on these 46 probe sets (Supplementary Table S1).

Modified DS index. The normalized gene expression files for the 23 dasatinib-treated breast cancer cell lines were used to identify differentially expressed genes between dasatinib-sensitive and dasatinib-resistant cells using two-sample unequal variance *t* tests. The top 142 probe sets that showed false discovery rates of <10% (false discovery rate, <0.1) were selected as the informative gene set (Supplementary Table S2).

Target index. This index was defined as the weighted sum of the expression of 19 distinct kinases that dasatinib binds to with high affinity (KD50, <1 nmol/L) *in vitro*. When multiple probe sets corresponded to one gene, the probe set with the highest median intensity and greatest interquartile range was selected as the representative probe set (Supplementary Table S3).

Statistical methods for developing response classifiers and assessing quality control

To develop the mDS index or the modified Src index, we explored three different classification algorithms

including weighted voting method, centroid method, and/or support vector machines (SVM).

Weighted voting. This predictor used the weighted voting algorithm (18) and it is similar to the one described in the publication of dasatinib sensitivity signature in cell lines (12). Initially, [$\mu_1(g)$, $\sigma_1(g)$] and [$\mu_2(g)$, $\sigma_2(g)$] were defined as the medians and SDs of the log expression levels of gene *g* for the cell line samples in class 1 (sensitive class) and class 2 (resistant class). The weight for gene *g* was defined as

$$a_g = [\mu_1(g) - \mu_2(g)] / [\sigma_1(g) + \sigma_2(g)];$$

The decision boundary for *g* was defined as

$$b_g = [\mu_1(g) + \mu_2(g)] / 2.$$

Whenever x_g denoted the normalized log expression value of gene *g* in a new sample, then the vote for gene *g* was $(x_g - b_g)$ and the weighted vote of gene *g* would be $v_g = a_g(x_g - b_g)$, with a positive value indicating a vote for class 1 and negative value indicating a vote for class 2. The votes of the selected genes were summed to obtain total votes. For class 1,

$$V_1 = \sum_{g=1}^n v_g * I_1,$$

in which $I_1 = 1$ if $v_g > 0$, otherwise $I_1 = 0$; for class 2,

$$V_2 = \sum_{g=1}^n v_g * I_2$$

in which $I_2 = 1$ if $v_g < 0$, otherwise $I_2 = 0$. The new sample would be assigned to the class receiving the higher number of absolute value of votes.

Centroid method. Centroids were defined as the median log-transformed expression values of the informative probe sets in the sensitive class or Src-transformed class (positive reference) and the resistant class or Src control class (negative reference). For a new sample, the similarity

Table 1. Performance of three different cell line–derived dasatinib sensitivity (mDS) predictors in independent cell lines treated with dasatinib

	Data set	ACC	Sensitivity	Specificity	PPV	NPV	AUC
Weighted Voting	Breast cancer (<i>n</i> = 12)	0.917	0.750	1.000	1.000	0.889	0.906
	Lung cancer (<i>n</i> = 23)	0.739	0.909	0.583	0.667	0.875	0.818
Centroid	Breast cancer (<i>n</i> = 12)	0.917	0.750	1.000	1.000	0.889	0.906
	Lung cancer (<i>n</i> = 23)	0.696	0.455	0.917	0.833	0.647	0.833
SVM	Breast cancer (<i>n</i> = 12)	0.667	0.000	1.000	NA	0.667	0.938
	Lung cancer (<i>n</i> = 23)	0.522	0.091	0.917	0.500	0.524	0.758

Abbreviations: ACC, overall accuracy; PPV, positive predictive; NPV, negative predictive value; AUC, area under the receiver operator characteristic curve. NA=not applicable. This model identified every cell line as a non-responder, thus PPV could not be calculated.

between its log-transformed gene signature and each reference centroid was calculated. Hoeffding's D statistic (19) was measured. It is a nonparametric measure of association like Spearman's rank but it takes into account the joint rankings of the two variables as previously published (20).

Hoeffding's D is robust against a wide variety of alternatives to independence, such as nonmonotonic relationships. The larger the value of D, the more dependent the two variables are. The difference between the two statistics D^+ and D^- was obtained, in which D^+ was Hoeffding's D relative to the positive centroid and D^- was relative to the negative centroid. The difference ($D^+ - D^-$) is the index for the new sample.

Support vector machine. A SVM performs classification by constructing an N -dimensional hyperplane that optimally separates the data into two categories. The points lying on the boundaries are called support vectors and the middle of the margin is the optimal separating hyperplane. Data points on the "wrong" side of the discriminant margin are weighted down to reduce their influence. The task can be formulated as a quadratic optimization problem (21, 22). The predictor based on SVM was developed using an R interface to the package *libsvm*.

To calculate the T index, the normalized, \log_2 -transformed gene expression values for each of the 19 distinct kinases targeted by dasatinib were weighed by the corresponding KD_{50} values (biochemical binding constants) from the *in vitro* experiments.(13) Individuals with the highest 25% of T index values were labeled as potential "responders" and therefore eligible for therapy. This 25% cutoff is arbitrary but it is based on the hypothesis that the higher the T index value, the greater the sensitivity to dasatinib.

After generating response classifiers using the methods described above, each classifier was evaluated to determine the incidence of predicted responders, the overlap for prediction of responding patients among the three indices, technical reproducibility, and variation incurred using differing tissue acquisition methods (FNA versus core needle biopsy) to select the most robust method for clinical use.

Results

Generation of mDS, Src, and T indices

The cell line-derived dasatinib sensitivity (mDS) index was developed as follows. Three different classification methods (weighted voting, centroid method, and SVM) were used to build prediction models in combination with the top-ranked probe sets from comparison of sensitive and resistant breast cancer cell lines ($n = 23$). The Monte Carlo complete cross-validation (3-fold, 500 times) was used to assess the prediction performance in the training set. The top three predictors that showed the nominally best performance in cross-validation were also tested on independent cell line validation data sets,

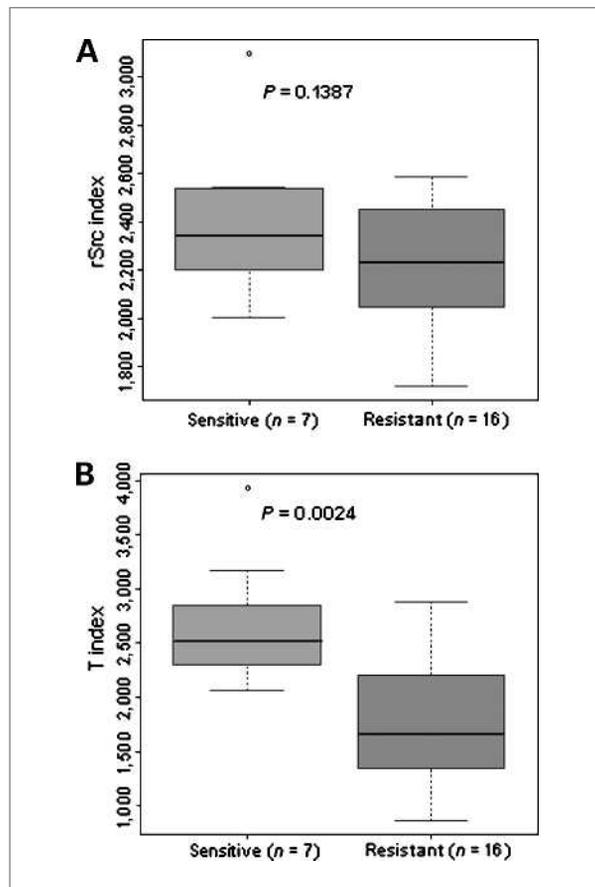


Figure 1. Predictive capability of rSrc index (A) and T index (B) in breast cancer cell lines with known sensitivity to dasatinib. Twenty-three breast cancer cell lines were categorized using the mean \log_{10} IC_{50} values (0.6 $\mu\text{mol/L}$ of dasatinib) to assign a sensitivity category to each cell line as previously published (13). Data from these cell lines were preprocessed with the Microarray Suite version 5.0 using global scaling as the normalization method. Predicted response scores were calculated using the rSrc and T indices as described in text. Differences in response score between responding and nonresponding cell lines were evaluated using the Student's t test.

which included 12 breast cancer and 23 lung cancer cell lines, each with known dasatinib sensitivity. The prediction results and performance metrics in this validation set are presented in Table 1. In general, better performance was seen in breast cancer compared with lung cancer cell lines and the weighted voting and centroid methods outperformed SVM that had high specificity but very low sensitivity in the validation cohort. Next, we examined the concordance of prediction results in three distinct sets of gene expression data generated in three different laboratories for the same 14 breast cancer cell lines. The genomic response prediction results were largely concordant for the same cell lines from different laboratories. The concordance correlation coefficients ranged from 0.676 to 0.912 (Supplementary Table S4). The weighted voting-based method was selected as the final cell line-derived mDS predictor for clinical testing.

Development of the Src pathway activity scores. First, we developed a Src pathway predictor that used the centroid method to measure the similarity between test samples and cell line centroids. This approach failed to show reproducibility in the replicate cell line and in the human data (the concordance correlation coefficient in replicate experiments was -0.13), and therefore, the method was revised to use the average expression of the 46 informative genes as measure of Src pathway expression, which yielded more robust results. This revised Src index (rSrc) was calculated for 133 breast cancer samples and the top 25th percentile for this reference population (score, $>2,124$) was used as cutoff to identify potentially dasatinib-sensitive patients. The rationale behind the selection of this threshold is that the higher the average expression of the Src-regulated genes, the greater the sensitivity to Src inhibition may be. When this rSrc pathway index was calculated for the 23 cell lines with known dasatinib sensitivity, the index showed no significant difference between the sensitive and resistant groups (Fig. 1A).

Development of the dasatinib T index. To calculate this index, the normalized, \log_2 -transformed gene expression values for each of the 19 kinases targeted by dasatinib were weighed by the corresponding KD50 values (i.e., binding constants). The cutoff value of 3,070 was derived from the reference population of 133 breast cancer cases and defines the lower limit of the target indices that belong to the top 25% of the overall index distribution (Supplementary Fig. S1). When the T index was tested

using the 23 cell lines with known dasatinib sensitivity, the sensitive cells had statistically significantly higher T scores than resistant cell lines (Fig. 1B).

Overlap of prediction results in a reference population of early-stage breast cancer

Genomic data from 133 breast cancer samples was used to determine the incidence of potential "responders" as predicted by the different genomic predictors and to assess overlap and concordance in predicted response by the different methods. The Src pathway activity and kinase T indices each identified 33 patients (25% by definition) as potential responders to dasatinib, whereas the cell line-derived response predictor index identified 13 (10%) patients as potential responders. Three patients (2%) were predicted to be sensitive by the all three predictors (Fig. 2). The prediction concordance between the mDS, rSrc, and T indices are shown in Table 2. The results indicate that these three conceptually different candidate predictors of dasatinib response identify largely distinct and nonoverlapping patient populations as potentially sensitive to this drug.

Technical reproducibility of prediction results in replicate human data and in different tissue samples from the same cancer

Gene expression data from 39 technical replicates, patient samples that were split and profiled twice, were used to assess the reproducibility of prediction results. The concordance correlation coefficients were 0.97, 0.95,

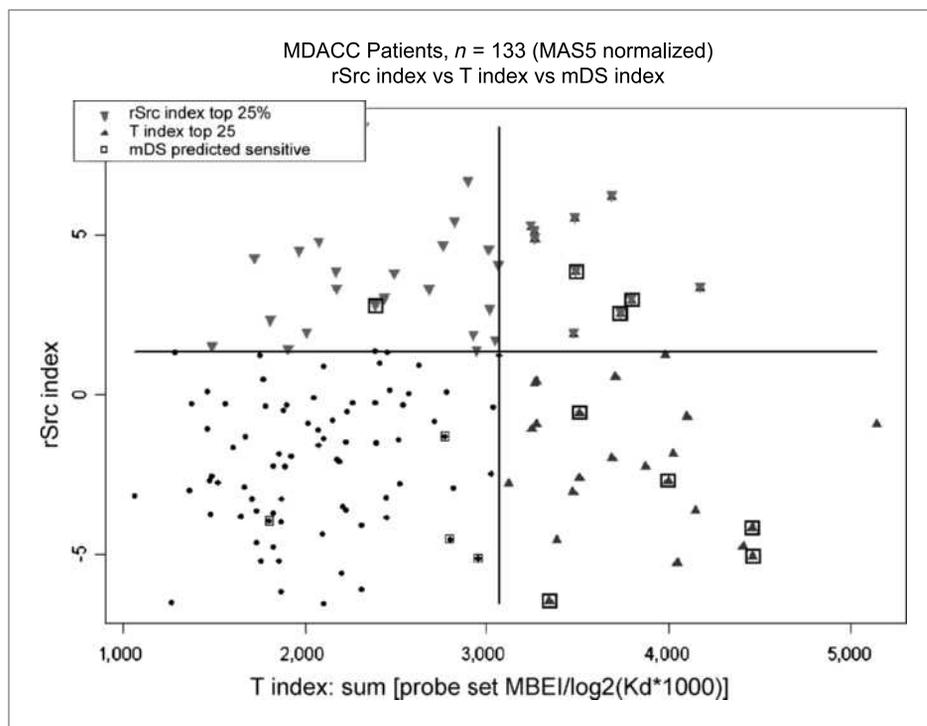


Figure 2. Overlap of response prediction for mDS index, rSrc index, and T index in human breast cancer data. A reference population of 133 nontreated breast cancer samples was used to determine the incidence of potential responders as predicted by the mDS index, the rSrc index, and the T index. All gene expression data were generated using Affymetrix HG-U133A chips.

Table 2. Overlap of predictors of dasatinib response in human breast cancer data

	T index	
	Sensitive	Resistant
mDS index		
Sensitive	8	5
Resistant	25	95
	rSrc index	
	Sensitive	Resistant
mDS index		
Sensitive	0	13
Resistant	33	87
	rSrc index	
	Sensitive	Resistant
T index		
Sensitive	2	31
Resistant	31	69

and 0.74 for the mDS, T index, and the rSrc index, respectively (Fig. 3).

To assess the effect of tissue sampling methods on the reproducibility of prediction results we examined paired samples of FNA and core needle biopsies from the same primary tumors ($n = 37$ pairs; Fig. 4). FNA specimens of cancer contain predominantly neoplastic cells (80–95%) and are generally devoid of stromal elements, other than a few infiltrating leukocytes (21). In contrast, core needle biopsies contain a variable amount of stromal components including fat cells, fibroblasts, and endothelial cells. Concordance correlation coefficients were modest ranging from 0.21 to 0.38, indicating that uniform biopsy method is critical for obtaining reproducible results. Improved normalization and data scaling methods needs to be developed to improve the concordance of genomic prediction results across different tissue sample types.

Discussion

Patient selection has always been a laudable goal of targeted therapy and early successes, such as trastuzumab for the treatment of human epidermal growth factor receptor 2–positive breast cancer and imatinib for c-KIT–expressing gastrointestinal stromal tumors, fostered the concept of tailored therapy to enhance response in clinical trials and reduce unnecessary patient exposure to inactive agents (23, 24). Unfortunately, this goal remains elusive for almost all other targeted agents (24). The need to discover response markers has become more pressing as clinical trials started to show tantalizing but low overall clinical benefit rates for most biologically targeted

agents. We previously suggested that it is possible to prospectively test candidate response markers in the clinic in a similar manner as we test new drugs in phase II studies to assess their potential clinical value as patient selection tools (9).

The current study describes the development of three conceptually different potential gene expression–based predictors of dasatinib response for clinical evaluation. For almost all drugs, one could rationally propose

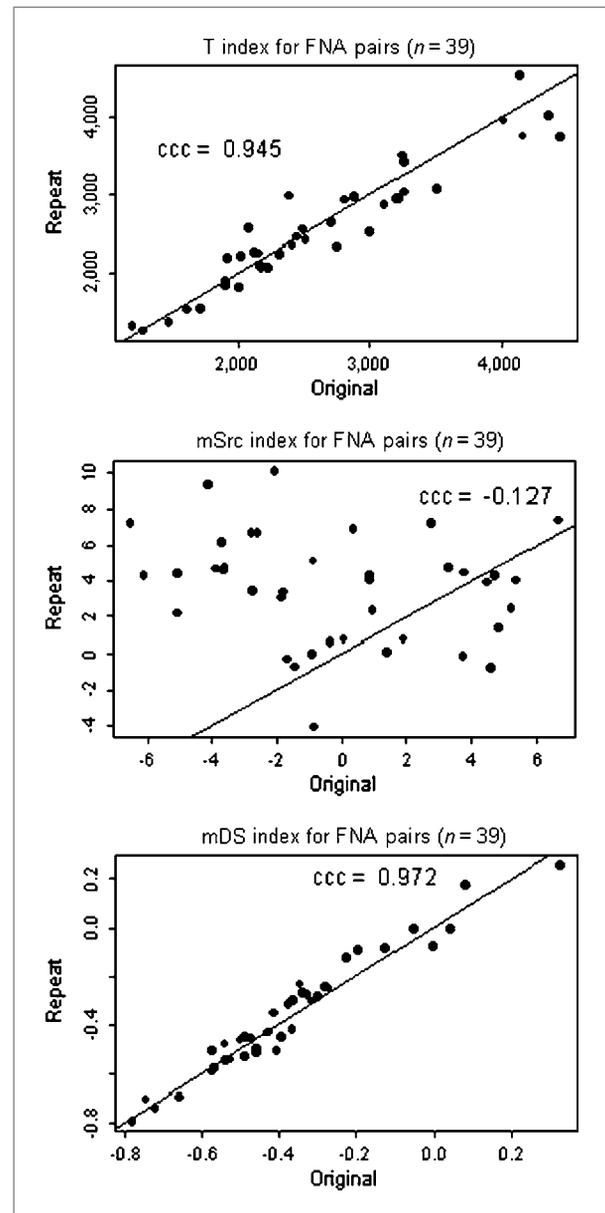


Figure 3. Technical reproducibility of mDS index, modified Src index (mSrc), and T index. To measure the reproducibility, 39 patient samples obtained by FNA were split and profiled twice either in different laboratories or several months apart in the same laboratory.

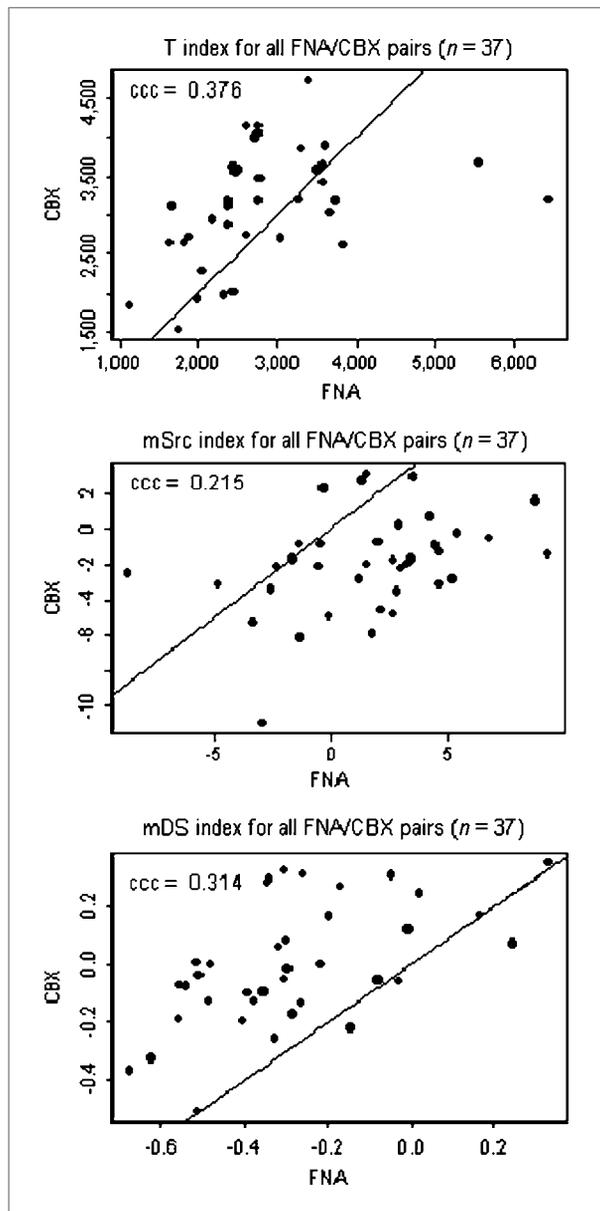


Figure 4. Concordance of predictors in tissue obtained by FNA or core needle biopsy (CBX). To determine the effects of sample acquisition methods (FNA versus core needle biopsy), 37 matched pairs of tissue samples obtained by FNA and core needle biopsy were analyzed.

candidate predictors based on the known or presumed mechanism of action of the drug and based on data from preclinical models. We applied this strategy to dasatinib and used empirical data from cell line experiments *in vitro*, information from mechanism of action (i.e., inhibition of multiple kinases), and biological insights into cancer biology (i.e., Src activation in breast cancer) to develop several different potential response predictors. Interestingly, marked differences were seen in the reproducibility and robustness of some predictors across cell lines and human samples. For example, a

Src pathway activity predictor using cell line centroids yielded results that had low reproducibility in replicate experiments and the predictions were unstable in human breast cancer samples. Data from human tissues did not correlate highly with any of the cell line centroids and therefore small changes in the data due to noise could influence class assignment.

We also show that there was substantial discrepancy in prediction results generated by the same prediction method on matching FNA and core needle biopsies of the same tumor. This is likely due to the different cellular composition of these tissues and the confounding effects of tumor stroma in core biopsies (25, 26). This is an important observation that suggests that clinical trials that assess biological markers should not use these sampling methods interchangeably without showing robustness of the predictor across tissue sampling methods. Hypothetically, both methods of sampling have their advantages. Biopsies obtained by FNA are usually less painful to patients and offer a higher percentage of cancer cells for response prediction. Although core needle biopsy can be more painful, it includes higher proportion of stroma; thus the interaction between tumor cell and nonmalignant surrounding tissue are both subject to analysis (27).

An interesting observation in this study was that the cell line-derived predictor, the Src pathway activity score, and the target expression index each identified distinct and only minimally overlapping patient populations as possibly sensitive to dasatinib. Considering the very different conceptual underpinnings of these predictors, this may not be surprising from a technical point of view. The true predictive values of the three candidate markers described in this study are unknown. As mentioned in the introduction, we have initiated a study that uses these predictors to select patients for dasatinib therapy (clinicaltrials.gov ID: NCT00780676); only individuals whose cancer is predicted to respond by one of these methods will receive dasatinib and early stopping rules apply to each marker group to stop accrual if less than expected clinical benefit is observed in the molecularly selected patient subset.

Disclosure of Potential Conflicts of Interest

G.N. Hortobagyi: advisor to Bristol-Myers Squibb. L. Pusztai: commercial research grant and honoraria from Speakers Bureau, Bristol-Myers Squibb. F. Huang: expert testimony, Bristol-Myers Squibb. No other potential conflicts of interest were disclosed.

Grant Support

American Society of Clinical Oncology Advanced Clinical Research Award and grants from the Breast Cancer Research Foundation (L. Pusztai).

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.

Received 07/31/2009; revised 01/25/2010; accepted 02/09/2010; published OnlineFirst 04/27/2010.

References

1. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
2. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–6.
3. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999–2009.
4. Yu K, Lee CH, Tan PH, et al. A molecular signature of the Nottingham prognostic index in breast cancer. *Cancer Res* 2004;64:2962–8.
5. Liu R, Wang X, Chen GY, et al. The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* 2007;356:217–26.
6. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817–26.
7. Ayers M, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 2004;22:2284–93.
8. Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 2006;24:4236–44.
9. Pusztai L, Anderson K, Hess KR. Pharmacogenomic predictor discovery in phase II clinical trials for breast cancer. *Clin Cancer Res* 2007;13:6080–6.
10. Fin RS BC, Ibrahim N, Strauss SC, et al. Phase II trial of dasatinib in triple negative breast cancer: results of study CA180059. 2008 San Antonio Breast Cancer Symposium. abstract 3118.
11. <http://clinicaltrials.gov/ct2/show/NCT00780676?term=dasatinib&recr=Open&rslt=Without&type=Intr&rank=16>.
12. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439:353–7.
13. Huang F, Reeves K, Han X, et al. Identification of candidate molecular markers predicting sensitivity in solid tumors to dasatinib: rationale for patient selection. *Cancer Res* 2007;67:2226–38.
14. Carter TA, Wodicka LM, Shah NP, et al. Inhibition of drug-resistant mutants of ABL, KIT, and EGF receptor kinases. *Proc Natl Acad Sci U S A* 2005;102:11011–6.
15. <http://bioinformatics.mdanderson.org/pubdata.html>.
16. <http://cancer.lbl.gov/breastcancer/data.php>.
17. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 2006;10:515–27.
18. Slonim DK, Tamayo P, Mesirov JP, Golub TR, Lander ES. Class prediction and discovery using gene expression data. In *Proceedings of the Fourth Annual International Conference on Computational Molecular Biology (RECOMB) 2000*;2000:263–72.
19. In: Hollander M, Wolfe DA, editors. *Nonparametric Statistical Methods*. 2nd ed New York: Wiley; 1999.
20. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer-Verlag; 2001.
21. Meyer D, Leisch F, Hornik K. The support vector machine under test. *Neurocomputing* 2003;55:169–86.
22. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. 2001. Software available from: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
23. Baselga J. Herceptin alone or in combination with chemotherapy in the treatment of HER2-positive metastatic breast cancer: pivotal trials. *Oncology* 2001;61 Suppl 2:14–21.
24. Le Tourneau C, Vidal L, Siu LL. Progress and challenges in the identification of biomarkers for EGFR and VEGFR targeting anticancer agents. *Drug Resist Updat* 2008;11:99–109.
25. Pusztai L, Ayers M, Stec J, et al. Gene expression profiles obtained from fine-needle aspirations of breast cancer reliably identify routine prognostic markers and reveal large-scale molecular differences between estrogen-negative and estrogen-positive tumors. *Clin Cancer Res* 2003;9:2406–15.
26. Symmans WF, Ayers M, Clark EA, et al. Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma. *Cancer* 2003;97:2960–71.
27. Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med* 2008;14:518–27.

Molecular Cancer Therapeutics

Development of Candidate Genomic Markers to Select Breast Cancer Patients for Dasatinib Therapy

Stacy Moulder, Kai Yan, Fei Huang, et al.

Mol Cancer Ther 2010;9:1120-1127. Published OnlineFirst April 27, 2010.

Updated version	Access the most recent version of this article at: doi: 10.1158/1535-7163.MCT-09-1117
Supplementary Material	Access the most recent supplemental material at: http://mct.aacrjournals.org/content/suppl/2010/04/26/1535-7163.MCT-09-1117.DC1

Cited articles	This article cites 20 articles, 8 of which you can access for free at: http://mct.aacrjournals.org/content/9/5/1120.full#ref-list-1
Citing articles	This article has been cited by 7 HighWire-hosted articles. Access the articles at: http://mct.aacrjournals.org/content/9/5/1120.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, use this link http://mct.aacrjournals.org/content/9/5/1120 . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.