Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival

Howard Y. Chang, Dmitry S. A. Nuyten, Julie B. Sneddon, Trevor Hastie, Robert Tibshirani, Therese Sorlie, Hongyue Dai, Yudong D. He, Laura J. van’t Veer, Harry Bartelink, Matt van de Rijn, Patrick O. Brown, and Marc J. van de Vijver

Contributed by Patrick O. Brown, January 5, 2005

Based on the hypothesis that features of the molecular program of normal wound healing might play an important role in cancer metastasis, we previously identified consistent features in the transcriptional response of normal fibroblasts to serum, and used this “wound-response signature” to reveal links between wound healing and cancer progression in a variety of common epithelial tumors. Here, in a consecutive series of 295 early breast cancer patients, we show that both overall survival and distant metastasis-free survival are markedly diminished in patients whose tumors expressed this wound-response signature compared to tumors that did not express this signature. A gene expression centroid of the wound-response signature provides a basis for prospectively assigning a prognostic score that can be scaled to suit different clinical purposes. The wound-response signature improves risk stratification independently of known clinico-pathologic risk factors and previously established prognostic signatures based on unsupervised hierarchical clustering (“molecular subtypes”) or supervised predictors of metastasis (“70-gene prognosis signature”).

response” (CSR) genes and their canonical expression pattern in fibroblasts activated with serum, the soluble fraction of clotted blood and an important initiator of wound healing in vivo. The CSR genes were chosen to minimize overlap with cell cycle genes, but instead appeared to represent other important processes in wound healing, such as matrix remodeling, cell motility, and angiogenesis, processes that are likely also to contribute to cancer invasion and metastasis. In several common epithelial tumors such as breast, lung, and gastric cancers, expression of the wound-response signature predicted poor overall survival and increased risk of metastasis (10). These initial findings demonstrate the promise of using hypothesis-driven gene expression signatures to provide insights from existing gene expression profiles of cancers. However, as in other methodologies, reproducibility and scales for interpretation need to be evaluated before this strategy can be generally adopted for biologic discovery and clinical use.

The best validation of a gene signature’s prognostic value is to test its ability to predict outcome in large independent data sets. Here we examine a database of 295 breast cancer patients from the Netherlands Cancer Institute that had previously been used to identify and validate a prognostic gene expression profile defined by a set of 70 genes (5, 9). We used this data set to test the reproducibility of the association between the wound-response signature and breast cancer progression, and to investigate how the information from diverse gene expression signatures identified by various means might be integrated both biologically and for clinical use.

Materials and Methods

Tumor Gene Expression Profiles. RNA isolation, labeling of complementary RNA, competitive hybridization of each tumor cRNA with pooled reference cRNA from all samples to 25,000 element oligonucleotide microarrays, and measurement of expression ratios were described (5). Detailed patient information has been described (9). Adjuvant setting was based on national guidelines or determined by participation in clinical trials at the time of diagnosis. Ten of the 151 patients who had lymph......

Data Analysis. Prognostic signatures. Genes on Stanford cDNA microarrays and Rosetta/NKI oligonucleotide microarrays were mapped between different platforms by using Unigene identifiers (build 158, release date January 18, 2003). This older build of Unigene was used to allow comparison with two published cross-platform analyses (10, 13). In the unsupervised analysis, 295 tumor samples were grouped by similarity of the expression pattern of the CSR genes by average linkage clustering by using the software CLUSTER (14); the gene expression values were centered by mean. The samples were segregated into two classes based on the first bifurcation of the clustering dendrogram; the two classes were identified as “activated” vs. “quiescent” by the predominant expression of the serum-induced and serum repressed CSR genes (10). Classification of the tumors as having a good prognosis signature or a poor prognosis signature based on the expression of 70 genes was as described (9). The five-class “intrinsic gene” signature was assigned by matching the expression value of the intrinsic genes in the NKI dataset to the nearest expression centroid of the five classes as described; samples that did not have correlation > 0.1 to any centroid were termed unclassified (13). A total of 509 probes representing 431 of 487 intrinsic genes were successfully identified in the NKI data set.

Survival analysis. Overall survival was defined by death from any cause. Distant metastasis-free probability (DMFP) was defined by a distant metastasis as a first recurrence event; data on all patients were censored on the date of the last follow-up visit, death from causes other than breast cancer, the recurrence of local or regional disease, or the development of a second primary cancer, including contralateral breast cancer. Kaplan–Meier survival curves were compared by the Cox–Mantel log-rank test in Winstat for Excel (R. Fitch Software, Staufen, Germany). Multivariate analysis by the Cox proportional hazard method was performed by using the software package SPSS 11.5 (SPSS, Chicago).

Scaling the wound signature. The patient data set was randomized into two halves, one for training and one for testing. The two half sets were matched for all known clinical parameters and risk factors (Table 2, which is published as supporting information on the PNAS web site). The serum-activated fibroblast centroid was as described (10). Pearson correlation of the expression values of CSR genes of tumor samples to the serum-activated fibroblast centroid results in a quantitative score reflecting the wound-response signature for each sample. The higher the correlation value, the more the sample resembles serum-activated fibroblasts (“activated” wound-response signature). A negative correlation value indicates the opposite behavior and higher expression of the “quiescent” wound-response signature. The threshold for the two classes can be moved up or down from zero depending on the clinical goal. Sensitivity and specificity for predicting metastasis as the first recurrence event was calculated for every threshold between −1 and +1 for the correlation score in 0.05 increments. The threshold value of negative 0.15 correlation gave 90% sensitivity for metastasis prediction in the training set, and had equivalent performance in the test set.

Decision tree analysis. To construct a decision tree, we considered all clinical risk factors and gene expression profiles by using the Cox proportional hazard model in SPSS, identified the dominant risk factor (most significant P value) to segregate patients, and reiterated the process on each subgroup until the patients or risk factors became exhausted. For gene expression signatures, we used the correlation value to each canonical centroid as a continuous variable to capture the possibility that different thresholds may be optimal in different subgroups. Because 60 patients with lymph node-negative disease in this series were used to train the 70-gene signature (5), performance of the decision tree incorporating the 70-gene signature was validated on the independent subset of patients with lymph node-positive disease. The threshold for the 70-gene signature has been reported (5); the threshold for the wound-response signature was chosen based on outcome data in the training set. Performance of the decision tree analysis was validated by equal performance in the randomized training and testing sets of patients. Support of the decision tree model by nonlinear multivariate analysis is described in Fig. 4, which is published as supporting information on the PNAS web site.

Results

Prognostic Value of a Wound Response Gene Expression Signature in Breast Cancer. To validate the prognostic value of the wound-response signature, we examined the expression of the core
A powerful prognostic indicator in breast cancer.

These results confirm that the wound-response signature is an independent predictor of metastasis and death (Fig. 6, which is published as supporting information on the PNAS web site). A Scalable Prognostic Score Based on the Wound-Response Signature. The correlation value to the serum-activated fibroblast centroid was modeled as a continuous variable; the hazard ratio per +1.0 correlation value is reported and represents the different risks at two ends of the spectrum. CI, confidence interval. The hazard ratio per +0.1 correlation value for death and metastasis are 1.27 (95% CI = 1.10–1.48) and 1.22 (95% CI = 1.06–1.40), respectively. Parameters found to be significant (P < 0.05) in the Cox proportional hazard model are shown in bold.

### Table 1. Multivariate analysis of risk factors for death and metastasis as the first recurrence event in early breast cancer

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Hazard ratio (95% CI)</th>
<th>P value</th>
<th>Hazard ratio (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wound response signature</td>
<td>11.18 (2.52–49.6)</td>
<td>0.001</td>
<td>7.25 (1.75–30.0)</td>
<td>0.006</td>
</tr>
<tr>
<td>Age, per decade</td>
<td>0.66 (0.45–0.95)</td>
<td>0.027</td>
<td>0.71 (0.50–1.00)</td>
<td>0.052</td>
</tr>
<tr>
<td>Diameter of tumor, per cm</td>
<td>1.02 (0.98–1.04)</td>
<td>0.270</td>
<td>1.03 (1.01–1.06)</td>
<td>0.008</td>
</tr>
<tr>
<td>Lymph node status, per positive node</td>
<td>1.05 (0.94–1.16)</td>
<td>0.371</td>
<td>1.10 (1.01–1.21)</td>
<td>0.035</td>
</tr>
<tr>
<td>Tumor grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 2 vs. 1</td>
<td>2.86 (0.96–8.5)</td>
<td>0.059</td>
<td>1.87 (0.86–4.07)</td>
<td>0.117</td>
</tr>
<tr>
<td>Grade 3 vs. 1</td>
<td>3.14 (1.02–9.6)</td>
<td>0.045</td>
<td>1.70 (0.74–3.90)</td>
<td>0.212</td>
</tr>
<tr>
<td>Vascular invasion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–3 vessels vs. 0 vessels</td>
<td>0.95 (0.35–2.52)</td>
<td>0.918</td>
<td>0.78 (0.32–1.87)</td>
<td>0.57</td>
</tr>
<tr>
<td>&gt;3 vessels vs. 0 vessels</td>
<td>1.88 (1.13–3.11)</td>
<td>0.014</td>
<td>1.65 (1.02–2.68)</td>
<td>0.043</td>
</tr>
<tr>
<td>Estrogen receptor status, positive vs. negative</td>
<td>0.49 (0.29–0.83)</td>
<td>0.008</td>
<td>0.82 (0.47–1.41)</td>
<td>0.468</td>
</tr>
<tr>
<td>Mastectomy, vs. breast conserving therapy</td>
<td>1.23 (0.76–2.01)</td>
<td>0.401</td>
<td>1.28 (0.80–2.04)</td>
<td>0.311</td>
</tr>
<tr>
<td>No adjuvant therapy, vs. chemo or hormonal therapy</td>
<td>1.42 (0.80–2.52)</td>
<td>0.291</td>
<td>2.24 (1.32–3.82)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

The correlation value to the serum-activated fibroblast centroid was modeled as a continuous variable; the hazard ratio per +1.0 correlation value is reported and represents the different risks at two ends of the spectrum. CI, confidence interval. The hazard ratio per +0.1 correlation value for death and metastasis are 1.27 (95% CI = 1.10–1.48) and 1.22 (95% CI = 1.06–1.40), respectively. Parameters found to be significant (P < 0.05) in the Cox proportional hazard model are shown in bold.

A Scalable Prognostic Score Based on the Wound-Response Signature. The previous analyses depended on stratifying tumors within a predefined group, relative to which each tumor is evaluated. To allow practical clinical use of the wound signature, we needed to develop a method to evaluate the presence and strength of this signature independently in any newly diagnosed cancer. Classification by hierarchical clustering provided a mathematically reasonable but biologically arbitrary threshold for assigning a cancer to one of two groups; it is preferable to treat the threshold as a parameter and quantify the confidence with which patients are assigned to each class. The threshold for calling a tumor sample wound-like could then be scaled to favor sensitivity or specificity, depending on the clinical scenario. For example, in a screening setting, it may be preferable to favor sensitivity, whereas a clinical test to determine therapies associated with high morbidity should have high specificity.

The expression pattern of CSR genes in serum-treated fibroblasts served as the prototype of the “activated” profile of the wound-response signature (10). Thus, we considered a strategy based on the correlation of the expression profile of CSR genes in each tumor sample to a vector representing the centroid of the differential expression in response to serum in cultured fibroblasts from 10 anatomic sites (10). The correlation value to the gene expression centroid of serum-activated fibroblast generates a continuous score that can be scaled. To evaluate the prognostic utility of the scalable wound signature, multivariate analysis of the wound signature with known clinical and pathologic risk factors for breast cancer progression showed that the wound signature is an independent predictor of metastasis and death and provides more prognostic information than any of the classic risk factors in the multivariate model (Table 1, hazard ratio of 7 and 11, respectively, P < 0.01). Because the expression pattern of CSR genes in serum-activated fibroblasts was discovered completely independently of tumor gene expression data or clinical outcome, the prognostic power of the serum-activated fibroblast centroid in breast cancer provides strong evidence of the biologic link between a wound response and cancer progression.

### Incorporating the Wound-Response Signature for Improved Clinical Decision Making

Because the wound-response signature provides improved risk prediction compared to traditional criteria, we examined the utility of a scalable wound signature in a clinical scenario: the decision to treat with adjuvant chemotherapy in early breast cancer. Approximately 30% of women with early breast cancer have clinically occult metastatic disease, and treatment with chemotherapy in addition to surgical excision and radiotherapy improves their outcomes (15). Uniform treatment of early breast cancer in women younger than 50 years of age with chemotherapy increases the 10-year survival from 71% to 78% (absolute benefit of 7%) for lymph node negative disease and...
from 42% to 53% (absolute benefit of 11%) for lymph node positive disease, but at the cost of exposing a large number of women who do not benefit (89–93% of all breast cancer patients) to the morbidities of chemotherapy. The absolute benefit of chemotherapy for older patients is even smaller (3.3% for node-negative and 2.7% for node-positive patients) (15). Clinical parameters, such as lymph node status, tumor size, and histological grade, can provide prognostic information (16) and are summarized in commonly used clinical guides for deciding whether to treat with chemotherapy such as the National Institute of Health (NIH) (17) or St. Gallen (18) consensus criteria. Nonetheless, risk stratification based on clinical parameters is far from perfect and, as a result, many women who are unlikely to benefit are treated with chemotherapy.

Because the presence of the wound-response signature in the primary tumor is associated with an increased risk of subsequent metastasis, we used a scalable wound-response signature to identify a subset of patients with a predicted risk of subsequent metastasis of <10%. Within this low-risk population, the expected absolute benefit from chemotherapy would be very small and the decision to forego chemotherapy might be justified. We used the serum-activated fibroblast centroid to assign a correlation score to each tumor in the data set. We set a threshold for the correlation score that was able to identify 90% of all patients with subsequent metastasis; this threshold was validated by first learning the threshold in half of the samples and showing an equivalent performance in the remaining half of the data set.

We then tested whether this scaled wound-response signature provided improved risk stratification compared to traditional clinical criteria. Indeed, patients who were assigned as high risk by the NIH or St. Gallen consensus criteria had heterogeneous outcomes, and within these sets of conventional “high risk” patients, the supervised wound response score was able to identify a subset of patients with a low risk of subsequent metastasis (Fig. 2A and B). A total of 185 of the patients represented in the NKI data set had never received adjuvant chemotherapy; the clinical outcomes of these patients allowed us to examine the appropriateness of decision for chemotherapy provided by the clinical guidelines or wound signature. As shown in Fig. 2C, the majority of patients who did not develop metastases in this series were stratified as high risk by the NIH or St. Gallen criteria, and according to these criteria would have been treated with chemotherapy that would not benefit them. The wound-response signature appropriately identified 90% of patients who developed metastases as the first recurrence (the end point of the supervised scaling), and at the same time would have spared 30% of women who did not develop metastasis from exposure to chemotherapy. These results illustrate the potential utility and improved risk stratification that might be achieved by scaling the wound-response signature to fit the prognostic goals in a clinical setting.

Integration of Diverse Gene Expression Signatures. How can we integrate the information from disparate prognostic signatures that have been identified for breast cancer to optimize risk stratification? We focused on three signatures that have been validated in independent studies and represent distinct analytic strategies. Perou et al. (3) used an unsupervised clustering strategy to identify subtypes of locally advanced breast tumors with pervasive differences in global gene expression patterns; the subtypes are thought to represent distinct biologic entities and were associated with different risk of metastasis (4, 13). At least five subtypes were distinguished, termed basal-like, ErbB2, luminal A, luminal B, and normal-like, and can be identified by the pattern of expression of a set of 500 “innocent genes.” In contrast, van’t Veer et al. (5) selected a 70-gene signature based on the association of expression each gene with the likelihood of metastasis within 5 years. The 70-gene signature was trained on a subset of the 295 patients studied in the present work and validated on the entire group of 295 patients (9). Finally, the wound-response signature was identified in a hypothesis-initiated approach that specifically tested the relationship between tumor progression and a gene expression program identified in an experimental model of a wound response (10). Importantly, these prognostic signatures are defined by expression patterns of distinct sets of genes with little overlap: only 22 genes are shared by two signatures (18 of these genes were shared between wound response and the intrinsic gene list), and no gene is present in all three signatures.

We used each of the three signatures to evaluate this series of 295 breast tumors and found that, despite their different derivations, the signatures gave overlapping and generally consistent predictions of outcomes (Fig. 3A). Many primary tumors from patients that developed subsequent metastasis and died expressed both the 70-gene poor prognosis signature and the wound-response signature. A small group of tumors with poor outcome were not identified as having a poor prognosis by the 70-gene signature but were highlighted by the wound-response signature (right side of Fig. 3A). Similarly, almost all of the
basal-like subgroup, so termed because they express markers characteristic of the basal epithelial cells in breast ducts, expressed the 70-gene poor prognosis signature and the activated wound-response signature \( (P < 0.001, \chi^2 \text{ test}) \). These results thus strongly support the idea that the basal-like tumors represent a distinct disease entity with an aggressive clinical course. However, apart from the basal-like subtype, many tumors had expression patterns that were indeterminate with respect to the subtypes as defined by the intrinsic genes; >100 of the 295 tumors could not be confidently assigned to any of the five subtypes defined by Perou et al. (3) and Sorlie et al. (4) (Fig. 7, which is published as supporting information on the PNAS web site). The limited ability to classify these cancers based on the available data may be due to the incomplete representation of genes that define the intrinsic gene list in this data set, or due to the fact that the genes that define this classification system were identified in locally advanced breast cancer samples and may not be optimal for classifying earlier stage cancers. In multivariate analysis combining (additively) known clinical risk factors with all three signatures, the 70-gene signature and wound-response signature provided independent and significant prognostic information, whereas the intrinsic genes did not (i.e., their prognostic information was subsumed by the other parameters in the model; Tables 3 and 4 and Supporting Text, which are published as supporting information on the PNAS web site).

As an alternative approach to considering information from multiple gene expression signatures for clinical risk stratification, we developed and tested a decision tree algorithm. At each node in the decision tree, we considered all clinical risk factors and gene expression profiles, selected the parameter and threshold that best segregated the patients with divergent outcomes, and reiterated the process on each resulting subgroup until the patients or risk factors were exhausted. We discovered that, in decision trees incorporating gene expression signatures, the 70-gene and wound-response signatures were sufficient to capture most of the prognostic information in only two steps (Fig. 3 B–D). Modeling of nonlinear interactions between the gene expression signatures and clinical risk factors independently yielded a similar conclusion (Fig. 4). For patients with early breast cancer and lymph node involvement, the key clinical decision is whether and how to treat with adjuvant chemotherapy. As reported (9), patients with the favorable 70-gene profile had \( \approx 90\% \) metastasis-free probability (group 0). Patients whose cancers had a poor-prognosis 70-gene profile, but lacked the activated wound-response signature, have a risk profile similar to the aggregated average baseline (group 1); patients whose cancers had both the activated wound-response signature and the 70-gene poor-prognosis signature had a risk of metastatic disease \( \approx 6.4\)-fold higher than did patients in group 0 (10-year DMFP of 89%, 78%, and 47%, respectively). Thus, the patients in group...
Chang et al.

0 might reasonably consider foregoing adjuvant chemotherapy, whereas the patients in group 2 have a risk profile more similar to patients with locally advanced disease and might be recommended for dose-dense or taxane-based adjuvant chemotherapy (19, 20). Together, these results illustrate that adding the wound-response signature to existing clinical, pathologic, and gene expression prognostic factors can significantly improve risk stratification and clinical decision making.

**Discussion**

We used an independent data set to confirm that a wound-response gene expression signature is a powerful predictor of clinical outcome in patients with early stage breast cancers. Together with our previous results on locally advanced breast, lung, and gastric cancer, these findings reinforce the concept that a gene expression program related to the physiological response to a wound is frequently activated in common human epithelial tumors, and confers increased risk of metastasis and cancer progression. In the future, methods that simplify the evaluation of this molecular signature will be required for allowing routine use of the wound-response signature in clinical decision making. Prospective studies are needed to determine whether treatment decisions based on the wound-response signature might benefit patients.

The molecular mechanisms that activate, sustain, and eventually shut off the wound-response-like gene expression program in tumors should be investigated. By delineating the risk for metastasis based on the wound-response signature, it is possible that these high-risk breast cancer patients might some day benefit from therapies that target the wound response.

We have examined approaches to parameterize the wound-response signature so that it can be evaluated in tumors individually to yield a quantitative score; the interpretation of the wound signature score can then be rationally directed to suit the clinical task. As a first step toward integrating diverse prognostic signatures, we examined the interactions and information provided by three independent methods for using global gene expression patterns to classify breast cancers and predict their clinical course: one that defined five molecular subtypes, one that was discovered by directly fitting to survival data, and one based on an *in vitro* model of a wound response. The different signatures each classified tumors into coherent and internally consistent groups, and where the signatures diverged, the combined information gave improved risk stratification compared to individual signatures. These results show that diverse analytic strategies are continuing to identify distinct molecular features that are related to poor prognosis in these tumors. Visualizing the connections between the different gene expression signatures suggests potential explanations for disparate clinical outcomes and sets the stage for directed experimentation. For example, the high level activation of the wound signatures in the basal-like subtype of breast cancers raises the possibility that basal epithelial cells in breast ducts have distinct roles in wound healing and may differentially regulate the CSR genes.

Direct approaches to building prognostic models from global gene expression data, by simply fitting the models to clinical outcome features are restricted to a palette of relatively simple models to avoid overfitting. An optimal model, reflecting the underlying pathogenic mechanisms, may be poorly represented of any of the models evaluated in these top-down supervised approaches and thus never discovered. Our results illustrate the potential advantages of a “bottom-up” approach building from gene expression signatures, developed to represent specific hypothesis about underlying pathogenic mechanism. Such an approach has the potential to outperform the top-down model-independent approaches in improving cancer stratification and clinical decision making. Moreover, this model-dependent bottom-up approach has the advantage of providing specific testable ideas about pathogenic mechanism and thereby potential targets for treatment.

This work was supported by National Institutes of Health Grants CA77097 and CA85129 (to P.O.B.) and AR050007 (to H.Y.C.), a National Science Foundation Predoctoral Fellowship (to J.B.S.), Dutch Cancer Society Grant NKB 2002-2575 (to M.v.d.V., D.N., L.v.t.V., and H.B.), and the Howard Hughes Medical Institute. P.O.B. is an Investigator of the Howard Hughes Medical Institute.

---


