

EDITORIALS



Sorting Out Breast-Cancer Gene Signatures

Joan Massagué, Ph.D.

The advent of technology for the analysis of human tissue samples for the activity of the entire genome initiated a quest for the molecular profiling of tumors. The aim of this quest was a better system for classifying cancers, a clarification of the origin of these diseases, a more accurate prognostic capacity than was previously available, and an improved ability to choose among possible therapies. Progress in the molecular profiling of solid tumors began with the identification of sets of genes whose expression can be used to classify breast cancers into five so-called intrinsic subtypes.¹ Soon thereafter, two independent comparisons of highly aggressive and less aggressive breast cancers yielded “poor-prognosis” gene signatures that are predictive of the clinical outcome.^{2,3} With time, the trickle of prognostic signatures for breast cancer grew to a small flood.⁴⁻⁸ In this issue of the *Journal*, Liu and colleagues report on the most recent entry into this series, a gene-expression signature that is associated with survival not only in breast cancer but also in lung cancer, prostate cancer, and medulloblastoma.⁹ These are exciting results, but the disconcerting reality is that the gene sets in the various studies are largely nonoverlapping, raising questions about their biologic significance and clinical implications.

One way to answer those questions is by identifying the aspect of tumor biology that the genetic signatures represent, depending on how each signature was obtained. The original intrinsic-subtype signature classifies breast cancers into histopathological entities, including a subtype with a negative estrogen-receptor status that originates in the basal epithelial layer of the mammary ducts, two subtypes with a positive estrogen-receptor status that were derived from luminal epithelial cells, a subtype that overexpresses the

human epidermal growth factor receptor-2 gene (*HER2*), and a subtype similar to normal tissue.¹ When analyzed to predict clinical outcomes, each of these subtypes was associated with a different survival rate, even though the classification system was not developed to predict clinical outcomes. In contrast, the two “poor-prognosis” gene signatures were derived from comparisons between highly aggressive and less aggressive primary tumors, in an effort to classify breast cancers according to the clinical outcome.^{2,3} Likewise, a “recurrence score” gene set that was prospectively selected on the basis of distant recurrence in patients who had node-negative breast cancer and who were treated with tamoxifen also correlates with the prognosis for recurrence.⁸

Recent studies have used concepts of cancer biology to predict clinical outcomes. A growing awareness of the important role of cancer-associated fibroblasts in tumor progression yielded the possibility that the gene-expression pattern of serum-stimulated fibroblasts in culture might be related to the gene-expression pattern of fibroblasts in wounded tissues, including tissue wounded by an invasive tumor. Indeed, patients who have breast tumors that express such a “wound-response” signature have a poor clinical outcome.⁵ Furthermore, reasoning that aggressive tumors are able to withstand hypoxic conditions, investigators defined a “hypoxia-response” signature that is associated with poor outcome in breast cancer and ovarian cancer.⁶ That there could be an association between an aggressive tumor and markers of a wounded stroma or of hypoxia makes sense. Such an association would suggest that cancer gene signatures reflect specific biologic traits of genetically unstable cancer cell populations that evolve into increasingly aggressive entities under environmental pressures to which

they adapt. The profusion of gene-expression signatures that are associated with clinical outcomes thus becomes less of a paradox.

Another tantalizing aspect of cancer biology led Liu and colleagues to seek a new prognostic signature.⁹ This group pursued the possibility that the growth of tumors is driven by a minority of pluripotent cells that are akin to cancer stem cells. The cue to this idea comes from hematologic cancers, whose origin in stem cells is well established.¹⁰ The presence of such stem cells in a tumor is of great interest, especially if these tumor-perpetuating minorities turn out to be the cell populations most resistant to therapy. In a previous study, this group used samples of pleural effusions obtained from patients with metastatic breast cancer to isolate rare subpopulations of tumorigenic breast-cancer cells that can regenerate the phenotypic diversity of the original tumor population when injected into immunodeficient mice.¹¹ The remainder of the cancer cells in these samples were nontumorigenic breast-cancer cells. On the basis of a comparison between the gene-expression profiles of tumorigenic breast-cancer cells and normal breast epithelial cells, Liu and colleagues now report an “invasiveness gene signature” (IGS) whose expression is associated with poor overall survival and metastasis-free survival in patients with various types of cancer. In patients with breast cancer, the IGS has the greatest prognostic capacity for those with moderately differentiated estrogen-receptor-positive tumors, a subgroup whose risk is otherwise difficult to assess.

These provocative findings raise a number of new questions. First, how can the gene-expression pattern of a minority stem-cell-like population be dominant enough to be detectable in bulk tumors? One interpretation of these findings is that aggressive tumors may be rich in cancer stem cells. This view, however, is at odds with the description of tumorigenic breast-cancer cells as a minority population. An alternative explanation is suggested by the fact that the IGS was obtained by comparing tumorigenic breast-cancer cells with normal breast epithelial cells, and not with nontumorigenic breast-cancer cells. The IGS may represent a signature of oncogenically transformed cells. By comparison, the changes in gene expression responsible for the phenotypic differences observed between tumorigenic breast-cancer cells and nontumorigenic breast-cancer cells may not

be extensive. A comparison of either tumorigenic breast-cancer cells or nontumorigenic breast-cancer cells with normal breast epithelial cells resulted in an overlap in the lists of genes of approximately 60%. The average hazard ratio (a statistical index of relative risk) for a poor outcome as predicted by the IGS was about 30% greater than that predicted by signatures derived from comparisons between nontumorigenic breast-cancer cells and normal breast epithelial cells. These observations suggest that, although tumorigenic breast-cancer cells contributed to the predictive power of the IGS, the signature may comprise several factors that are shared by tumorigenic and nontumorigenic breast-cancer cells.

If the IGS reflects some general aspect of the transformed state, what might that aspect be? The 186 genes that make up the IGS govern broad biologic activities, making it difficult to ascertain the biologic basis of the enhanced tumorigenicity. Although the IGS was named an “invasiveness” signature, no evidence is provided that it actually mediates invasion. Invasion is a concrete activity of tumor cells that involves many known genes that are not particularly prominent among those included in the IGS. The IGS includes genes involved in two signaling pathways that are activated in many aggressive tumors — the I κ B/NF κ B pathway in response to proinflammatory cytokines, and the RAS/MAPK pathway in response to cell-proliferation factors.

A classification of tumors directly based on the gene signatures of RAS and other deregulated pathways was recently reported.⁴ These “oncogenic pathway” signatures permit risk stratification in several types of cancer. The joint deregulation of the Wnt/ β -catenin pathway (which is typically active in stem or progenitor cells) and the Src pathway (a “proinvasive” pathway) carried the worst prognosis for patients with lung adenocarcinomas, breast-cancer tumors, or ovarian-cancer tumors. The presence of certain hyperactive pathways in a tumor suggests that the survival of its cells depends on these particular pathways, which in turn points to specific therapeutic opportunities.⁴ Another approach to the identification of relevant therapeutic targets is based on screening for genes associated with the ability of human tumor cells to form metastases in mice. This approach recently yielded a breast-cancer “lung metastasis” signature that is also linked to a poor clinical outcome.⁷

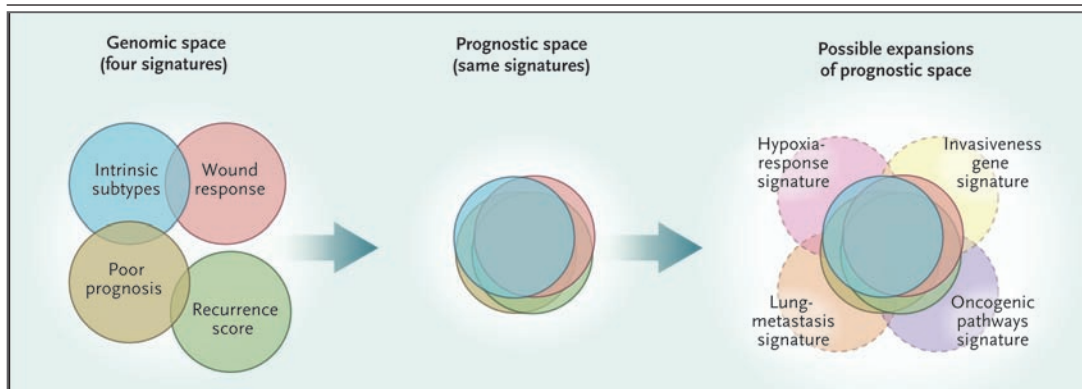


Figure 1. Four Gene-Expression Signatures.

The four gene-expression signatures shown, although derived by different means and containing largely nonoverlapping genes,^{1,2,5,8} show extensive overlap in predicting the clinical outcome for individual patients in the same cohort.¹² New gene-expression signatures that are based on functional criteria^{4,6,7,9} may provide complementary information that will help to refine a patient's prognosis and inform the therapeutic choices. Genomic space denotes the entire set of genes encoded in the human genome and therefore occupies a portion of the genomic space. Prognostic space denotes the range of prognostic possibilities covered by a particular prognostic indicator, in this case an indicator characterized by the altered expression of a particular set of genes.

Could different signatures be combined to provide a robust and accurate tool that would be useful in clinical practice? Several gene signatures — the intrinsic-subtype, poor-prognosis, wound-response, and recurrence-score signatures — were recently compared with this question in mind.¹² The signatures showed a significant agreement in the outcome predicted for the same patients. A model combining three of these signatures did not perform better than did each of the signatures separately. Thus, these four gene signatures may reflect a common set of phenotypic traits, with each trait defined by a set of gene-expression events. These signatures may be largely different from one another according to gene identity, but they occupy overlapping prognostic space — the range of prognostic possibilities covered by a particular prognostic indicator (Fig. 1). In other words, they may be regarded as different pictures of the same beast.

The power of combined signatures might be improved with the use of decision-tree analysis methods⁵ or the integration of new signatures (Fig. 1). For example, the performance of the IGS was better in predicting the clinical outcome when it was combined with the wound-response signature.⁹ Much work remains to be done to settle the questions discussed here. Cancer gene-expression signatures are beginning to be tested

in the clinic. Predictions of a good prognosis have been proposed as criteria for limiting the use of unnecessary adjuvant therapy. It remains to be seen whether such criteria would provide a sufficient incentive for the widespread use of prognostic signatures. Further incentives may be on the horizon: as pharmacologic inhibitors for specific pathways become available, the signatures that define tumors according to their vital pathways may provide crucial guidance for designing drug combinations of choice.

Dr. Massagué reports receiving consulting fees from Acceleron Pharma. No other potential conflict of interest relevant to this article was reported.

From the Cancer Biology and Genetics Program and Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York.

1. Sørlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418-23.
2. 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.
3. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;365:671-9.
4. 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.
5. Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci U S A* 2005;102:3738-43.

6. Chi JT, Wang Z, Nuyten DS, et al. Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med* 2006;3:e47.
7. Minn AJ, Gupta GP, Siegel PM, et al. Genes that mediate breast cancer metastasis to lung. *Nature* 2005;436:518-24.
8. 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.
9. 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.
10. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature* 2001;414:105-11.
11. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 2003;100:3983-8. [Erratum, *Proc Natl Acad Sci U S A* 2003;100:6890.]
12. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 2006;355:560-9.

Copyright © 2007 Massachusetts Medical Society.

The Mammogram That Cried Wolfe

Karla Kerlikowske, M.D.

That mammographic density is an important risk factor for breast cancer was first recognized by Wolfe in the 1970s. His pioneering observation has since been confirmed in more than 42 studies, the vast majority of which have shown an association between increased mammographic density and the risk of breast cancer.¹ Women in the highest quartile of mammographic density have a risk of breast cancer that is approximately four to six times as high as that among women of similar age who are in the lowest quartile. Only two other factors increase the risk of breast cancer more than mammographic density: age and mutations in the breast cancer-susceptibility genes *BRCA1* and *BRCA2*.

Mammographic density is a function of the abundance of epithelial and connective tissue in the breast, but a cancer and these normal tissues can have a similar radiographic attenuation, which can make both appear radiodense or white on a mammogram. By contrast, fat is radiolucent or dark on a mammogram.^{2,3} Therefore, it is possible that the risk associated with mammographic density is due to a masking effect — extensive breast density can hide a cancer.

In this issue of the *Journal*, Boyd et al.⁴ suggest that a masking effect is likely in the short term after mammography among women with density in 75% or more of the breast, as measured by qualitative or semiquantitative methods. They calculated the odds of screen-detected breast cancer (defined as a breast cancer detected at the time of screening mammography) and of breast cancer detected by methods other than screening (defined as a breast cancer that was detected within 12 months after a negative screening examination) in relation to the extent of mammographic density. They found that the odds ratio of screen-

detected breast cancer was 3.5 in women with extensive density as compared with women who had density in less than 10% of the breast. The masking effect greatly increased the odds of a cancer detected by nonscreening methods in women with extensive density as compared with those with density in less than 10% of the breast — odds ratio, 17.8.

The results of the study by Boyd et al. are similar to those of the Breast Cancer Surveillance Consortium, which reported that the rate of screen-detected breast cancer (defined as breast cancer detected within 12 months after a positive screening examination) is 2.5 times as high in women 40 to 49 years of age whose mammographic-density category was extremely dense as in women in the category called “almost entirely fat,” according to the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) (Table 1). The rate of breast cancer not detected by screening mammography (defined as breast cancer detected within 12 months of a negative screening examination) is 15 times as high in women 40 to 49 years of age with a BI-RADS density of extremely dense as in women whose mammogram is in the almost-entirely-fat category. Rates of cancer not detected by screening mammography among women with extensive breast density increase with age and are highest in women 60 to 69 years of age. A large proportion of women 40 to 49 years of age who receive a diagnosis of breast cancer have extensive mammographic density, but the absolute risk of cancer among women in this age group who have extensive density is relatively low — even lower than that among older women with extensive density (Table 1).

Boyd et al. also report an association between