Morgan, D.O., and Shokat, K.M. (1999). Chem. Biol. *6*, 671–678.

Papadopoulos, P., Ridge, S.A., Boucher, C.A., Stocking, C., and Wiedemann, L.M. (1995). Cancer Res. *55*, 34–38.

Pluk, H., Dorey, K., and Superti-Furga, G. (2002). Cell *108*, 247–259.

Sawyers, C.L. (1999). N. Engl. J. Med. *340*, 1330–1340.

Sawyers, C.L. (2001). Science *294*, 1834.

Schindler, T., Bornmann, W., Pellicena, P., Miller, W.T., Clarkson, B., and Kuriyan. J. (2000). Science *289*, 1938–1942.

Sicheri, F., Moarefi, I., and Kuriyan, J. (1997). Nature *385*, 602–609.

Xu, W., Harrison, S.C., and Eck, M.J. (1997). Nature *385*, 595–602.

Zhao, X., Ghaffari, S., Lodish, H., Malashkevich, V.N., and Kim, P.S. (2002). Nat. Struct. Biol. *9*, 117–120.

Functional genomics and the breast cancer problem

The clinical treatment of primary breast cancers has been greatly complicated by the inability to accurately predict which tumors will eventually become invasive and metastatic and which will remain localized and indolent. Lacking the ability to discriminate between these two classes of breast cancer patients, oncologists often apply aggressive adjuvant therapy to women in both groups. However, the use of functional genomics analysis has now made it possible to assemble a set of gene markers, the expression of which enables one to predict, with reasonably high accuracy, whether or not the patient will relapse or remain tumor-free five years after initial diagnosis and treatment.

The application of functional genomics to the analysis of breast cancer samples holds great promise in producing a substantial advance in breast cancer prognostication. In a recent issue of *Nature,* van't Veer et al. describe their ability to determine with substantial accuracy whether a young node-negative patient will or will not show progressive disease within five years time (van 't Veer et al., 2002).

This work highlights a painful disparity in current clinical practice: oncologists are able to detect tumors of ever smaller size (current avg.: ~1.5 cm diameter) in the breast, but their ability to exploit the diagnostic parameters associated with these relatively small tumors in order to predict eventual outcome has lagged far behind.

The oncologic territory has changed dramatically over the past two decades. With improving public awareness and detection techniques, as many as 60%–70% of breast cancer patients currently present clinically at an early stage while still being node-negative. That is, the axillary lymph nodes, which drain much of the mammary gland, are found to be free of detectable cancer cells through light microscopy.

Since the 10-year recurrence rate in node-positive patients is approximately 70%, the standard practice is to administer systemic adjuvant therapy to all patients in this group. This therapy involves anti-estrogen treatment (in the case of estrogen receptor-positive tumors), Herceptin (in the case of HER2/*neu*-positive tumors), and often cytotoxic drugs.

The treatment decision for node-negative patients, however, is not as straightforward. Approximately 70%–80% of the node-negative group will survive breast cancer without additional treatment beyond surgical resection of the initially detected growth**.** Hence, many argue that adjuvant systemic therapy, which can have significant toxicities, should not be undertaken routinely, since one risks harming many for the benefit of a few (Harris et al., 2000; Bland et al., 1998).

The dramatic shift in the clinical status of women at the time of initial presentation in the clinic now creates a substantial and still-to-be-solved problem, because many of the standard pathologic prognostic factors currently used were developed in an earlier era when a majority of the patients presented with larger (>2 cm diameter) tumors and were node-positive.

Among the many studies that have examined prognostic factors for breast cancer, only a few have specifically focused on node-negative patients (Mirza et al., 2002). Moreover, standard prognostic factors such as tumor size, histological type, grade of differentiation, and hormone receptor status were found to have limited value in predicting which node-negative patients would relapse. As a result, physicians have very few tools available for determining which node-negative patients should receive adjuvant therapy.

The landscape has also changed in yet another respect. Some of the traditional markers (e.g., estrogen receptor, ER) have been useful not only to predict outcome but also to inform therapy, since anti-ER agents such as tamoxifen have been available and found useful for improving short-term survival. Now, with the advent of novel therapies (e.g., Herceptin), other markers, such as Her2/*neu* expression, gain in importance because of their role in conferring responsiveness to these newly developed agents.

This complex picture reveals the disease of breast cancer to be a rapidly moving target, where the perceived relevance of widely used diagnostic and prognostic parameters may change dramatically over a period of several years.

The mysteries of breast cancer epidemiology

The number of diagnosed breast cancers in the Western world has increased substantially over the past half century. For example, in 1990, an estimated 150,000 new cases were reported in the United States while six years later, the number had increased to 185,000. During the same time period, estimated deaths due to breast cancer remained constant (~44,000) (Harris et al., 2000). However, there is some hope on the horizon: over the past several years, age-adjusted mortality of breast cancer has actually declined slightly (~1.7% per

year). The simplest interpretation of these statistics is that breast cancer is reaching epidemic proportions, but that advances in modern medicine have resulted in an increasing ability to cure disease. Accordingly, these two countervailing trends cancel each other almost precisely, leaving the mortality rate essentially unchanged for the past two generations.

There is, however, a far different interpretation of these facts, and it sheds less flattering light on the practice of medicine. Some would argue that the real incidence of breast cancer has changed only slightly over the past half century, and that the increased number of cases really represents a diagnostic bias. Thus, currently available diagnostic techniques may be detecting cancers in women that previously would have eluded detection and would have remained indolent for the lifetime of the woman, never emerging to the level of clinical visibility.

This perspective, in its extreme and most fatalistic form, assigns breast cancers to two categories—those that will remain non-life-threatening no matter how they are treated, and those that will progress, sooner or later, to a lethal outcome, once again no matter how they are treated. If this interpretation were sustained, then cynics might opine that a substantial proportion of breast cancer survivors have survived a form of the disease that in earlier days would have remained undetected and non-threatening during the lifetimes of these women. (A similar argument can also be made for prostate cancer.)

The real truth probably lies somewhere in between. For example, an ever-increasing average human life span might allow slowly growing, indolent tumors to progress to a clinically significant size and state of malignancy. Moreover, it is likely that there has been a real and significant increase in breast cancer incidence, and this increase remains even after the effects of diagnostic bias have been factored in. The most likely cause of this increase is due to the hormonal environment of Western women, which has changed profoundly over the past century. The point is made most starkly by referring to the fact that an 18- or 19-year-old woman in modern America has gone through as many menses as her great-grandmother experienced in a lifetime. This astounding conclusion flows directly from the increasingly early age of menarche, from delayed child-bearing (thereby forgoing the effects that both pregnancy and lactation have on suppressing menses), and from a menopause that is delayed by years compared with historical averages. These factors, together with women missing the protective effects that early and repeated childbearing exert in reducing breast cancer risk, should persuade even the most hardened cynics that the real incidence of breast cancer has indeed grown.

The oncologist's dilemma

The fact that the majority of breast cancers diagnosed these days are unlikely to lead to fatal outcomes, even without aggressive intervention, has been of little consolation to oncologists, who have been unable to determine with certainty which tumors will remain indolent and which, of an identical histology, will progress sooner or later to a highly malignant, life-threatening state. Because diagnostic bias has been increasing the numbers of breast cancers detected each year, the oncologist's dilemma worsens incrementally, since the proportion of diagnosed tumors that truly justify aggressive intervention decreases year after year. Being unable to distinguish truly worrisome tumors from those that will remain indolent, oncologists must treat all lesions equally aggressively. Hence, the number of breast cancer patients exposed unnecessarily to aggressive surgery and systemic adjuvant therapy (chemotherapy and axillary radiation) grows with each passing year.

Hope for successfully dealing with this dilemma has come from an increasingly diverse array of breast cancer markers. Ideally, the best of these markers should be proteins that play important causative roles in the tumor cell proliferation and thus cancer progression. The proteins commonly used in breast cancer prognostication include Her-2/Neu, p53, estrogen and progesterone receptors, and cell proliferation markers such as Ki-67 and cyclin D1. Historically, each of these markers emerged from detailed molecular and biochemical analyses of breast cancer cell lines and tumor samples.

Even when these markers are used together with tumor histopathology to determine the prognosis of a node-negative patient, the result is probabilistic and, as before, the eventual clinical course far from being a certainty. It remains unclear whether this ensemble of markers is intrinsically limited in its predictive powers, or whether the proper studies using all these diagnostic parameters together on a large patient population have yet to be done.

Enter functional genomics

Motivated by the shortcomings of current prognostic markers, van't Veer et al. screened the mRNA expression patterns of a set of 25,000 genes in tumor samples from node-negative breast cancer patients using an oligonucleotide microarray. Tumor-to-tumor comparisons revealed that 5000 genes varied in expression level by more than 2-fold. The great majority of these genes were found to have little value in predicting the known five-year survivals and recurrence rates of 78 node-negative patients, leaving a subset of 231 genes as potentially useful prognostic markers. A minimum set of 70 genes out of the 231, when analyzed with the proper algorithm, enabled these researchers to predict with 83% accuracy (i.e., 65/78 tumors) which of the node-negative patients registered in their data set would develop distant metastasis during a five year follow-up period and which would remain disease-free over this interval.

Provocatively, none of the clinically well-established prognostic marker proteins (i.e., members of the short list cited above) was encoded by a gene in the subset of 70 marker genes found to have prognostic utility in this study. This discrepancy could be explained by the fact that for many genes, there is little correlation between the abundance of its mRNA transcript with steady-state levels of encoded protein. Alternatively, the traditionally used markers may simply have far less predictive value than the 70 uncovered through this functional genomics strategy.

A logical weakness in retrospective studies like this one derives from the fact that a data set is used to generate a prognostic strategy, and the utility of the resulting strategy is then measured by examining its predictive powers on the same data set. However, when this functional genomics strategy was tested on a group of 19 patients who were not members of the original cohort, it worked well, yielding correct predictions in 17 of 19 patients examined. It remains to be seen whether application of this strategy to a large cohort of patients studied retrospectively, or to a large cohort studied

prospectively, will reveal predictive powers that are equally impressive. Moreover, 10- and 15-year outcomes will eventually prove to be more relevant, given the frequently delayed relapses of breast cancers after this long period of apparently tumor-free life.

Hope and problems for the future

Will we eventually be able to improve upon what van't Veer and colleagues accomplished in this study? When starting with a cohort of 25,000 genes, as was done in this paper, it was only possible to predict the actual outcome accurately in 83% of the patients. Does this represent an upper limit of success or could one achieve almost 100% accurate predictions with even more genes in the starting cohort (there being 6000–10,000 that were absent from their array)? An obvious source of improvement might come from dissection of the carcinoma cells from the stromal cells with which they are comingled in virtually all breast carcinomas. Many of the genes analyzed here are clearly expressed in one or the other cell type, and the confounding effects of stromal gene expression, contributed by larger or smaller stromal cell populations in different tumors, are hard to gauge in the absence of such dissection.

An absolute upper limit to predictive power might well be due to a subset of tumors that do not actually express the predictive information in the primary tumor sample. Problems may also arise from heterogeneity of tissue populations in primary tumor populations; thus, the sampled area may not represent the most aggressive portion of the tumor. Moreover, patient's exposure to unpredictable environmental factors (radiation, hormones, etc.) subsequent to tumor sampling may modify the behavior of the occult residual disease, both at local and distant sites.

Even with larger tumor samples and more powerful expression array analyses, a weighty problem may continue to dog those intent on generating more accurate breast cancer prognoses. Use of all these diagnostic assays, involving histopathology, ELISA assays, or gene expression arrays or combinations of two or three approaches, assume implicitly that individual breast cancers follow a small number of alternative genetic paths during the course of multi-step tumor progression. These paths, defined by the sequences of genetic and epigenetic alterations acquired by tumor cells during the development of malignancy, would seem to preordain the ultimate behavior of the end products of tumor progression.

The unsettling prospect is that breast cancer pathogenesis, and by extension the pathogenesis of many other human tumor types, is a far more random, disordered process, and that there are dozens if not hundreds of distinct routes taken by human breast cancers en route to full-fledged malignancy. If this were the case, then the predictive powers of diagnostic assays will forever be limited by the involvement of small but significant numbers of tumors that take their own, unique paths, each following the beat of a quite different drummer.

Some of the initial functional genomics analyses of tumors have demonstrated the power of this technology to stratify tumors into different subclasses. However, these proof-of-principle studies did not correlate their results with the clinical outcomes of the patients under study (Bittner et al., 2000; Perou et al., 2000). The study by van't Veer et al., together with another study in PNAS by Sorlie et. al., are among the first in the breast cancer field linking expression profiles to clinical outcome data (van't Veer et al., 2001; Sorlie et al., 2001). Similarly, two groups were recently successful in defining two subsets of diffuse large B-cell lymphoma (DLBL) patients having dramatically different outcomes that could not be predicted by conventional classification methods (Alizadeh et al., 2000; Scherf et al., 2000). With these reports on breast carcinomas and DLBL in hand, it is already obvious that a combination of clinical outcome data with functional genomics will yield a powerful approach for defining sets of genes whose expression will predict outcome and direct therapy far more effectively than current tumor markers allow.

Tan A. Ince[1,2] and
Robert A. Weinberg[2,3,4]

[1]Department of Pathology
Brigham and Women's Hospital
Harvard Medical School
Boston, Massachusetts 02115
[2]Whitehead Institute for Biomedical Research
Cambridge, Massachusetts 02142
[3]Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
[4]E-mail: weinberg@wi.mit.edu

Selected reading

Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X., et al. (2000). Nature *403*, 503–511.

Bland, K.I., Menck, H.R., Scott-Conner, C.E.H., Morrow, M., Winchester, D.J., and Winchester, D.P. (1998). Cancer *83*, 1262–1273.

Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., et al. (2000). Nature *406*, 536–540.

Harris, J.R., Lippman, M.E., Morrow, M., and Osborne, C.K. (2000). Diseases of the Breast (Philadelphia: Lippincott, Williams & Wilkins).

Mirza, A.N., Mirza, N.Q., Vlastos, G., and Singletary, S.E. (2002). Ann. Surg. *235*, 10–26.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Nature *406*, 747–752.

Scherf, U., Ross, D.T., Waltham, M., Smith, L.H., Lee, J.K., Tanabe, L., Kohn, K.W., Reinhold, W.C., Myers, T.G., Andrews, D.T., et al. (2000). Nat. Genet. *24*, 236–244.

Sorlie, T., Perou, C.M., Tibshirani R,. Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, et al. (2001). Proc. Natl. Acad. Sci. USA *98*, 10869–10874.

van't Veer, L.J., Dai, H., Van de Vijver, M.J., He, Y.D., Hart, A.A.M., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., et al. (2002). Nature *415*, 530–536.