about the book…

Pharmacogenetics is becoming increasingly relevant in the diagnosis, treatment, and recovery of cancer patients. A major problem facing oncologists is the outstanding varied efficacy of treatment. Promising advances in pharmacogenetics have allowed the development of effective agents which will enable personalized cancer chemotherapy to become routine for the clinical practice.

Written by experts in the field and combining information that is unable to be found in a single source volume, Pharmacogenetics of Breast Cancer:

- Combines a complete overview of pharmacogenetics and how it relates to breast oncology for diagnosis, treatment, and the recovery of patients using an individualized therapy model
- Is the first single source reference demonstrating the application of pharmacogenetics in the care of breast cancer patients
- Enables physicians a coherent interpretation of the emerging science of pharmacogenetics, aiding them to incorporate individual therapies in their own practice
- Gives practical guidance on various forms of specimen collection, tissue selection, and handling procedures

about the editor…

BRIAN LEYLAND-JONES is Associate Vice President and Director, Emory Winship Cancer Institute, Emory University, Atlanta, Georgia. Dr. Leyland-Jones’ main research interests are pharmacodynamics, pharmacokinetics, and pharmacogenetics in oncological clinical trials; translation of preclinical models into the clinic; biomarker endpoints in Phase I/II clinical trials; and screening and mechanistic studies of novel targeted and chemotherapeutic anti-cancer agents. He has authored more than 125 peer-reviewed articles and book contributions, 150 abstracts, and 29 patents.
TRANSLATIONAL MEDICINE SERIES

1. Prostate Cancer: Translational and Emerging Therapies, edited by Nancy A. Dawson and W. Kevin Kelly
2. Breast Cancer: Translational Therapeutic Strategies, edited by Gary Lyman and Harold Burstein
4. Multiple Myeloma: Translational and Emerging Therapies, edited by Kenneth C. Anderson and Irene Ghobrial
5. Cancer Supportive Care: Advances in Therapeutic Strategies, edited by Gary H. Lyman and Jeffrey Crawford
6. Cancer Vaccines: Challenges and Opportunities in Translation, edited by Adrian Bot and Mihail Obrocea
7. Pharmacogenetics of Breast Cancer: Towards the Individualization of Therapy, edited by Brian Leyland-Jones
Preface

I write this preface on Christmas Day, 2007, in Costa Rica. Christmas is always a special time for me, since my parents were married on Christmas Day sixty nine years ago. It is almost unimaginable to think how far we have progressed since 1938. My parents were married in the year before the Second World War. The nitrogen mustards, from which arose the first chemotherapies, were developed during the war. I was fortunate during my training at Sloan-Kettering to work in the laboratories of Fred Phillips, Joe Burchenal, and Jack Fox, which conducted some of the pivotal work on these compounds. Indeed, the first task force for childhood leukemia was set up in Washington in the year of my birth. However, we all have to remember how slowly treatments are disseminated to the community; I can remember, as a child, my next-door neighbor having a leg amputated for metastatic lung cancer since no other treatments were accessible locally!

The rest is history, but again it is difficult to believe that the Nobel Prize for the discovery of the cellular origin of retroviral oncogenes was awarded to Bishop and Varmus only in 1989.

Now we face the challenge of too few patients eligible for clinical trials to adequately test the myriad of targeted therapies available as well as the availability of high-quality starting material with associated clinical and phenotypic data. Various platforms are available for the identification of biomarkers: early technologies have focused on reverse transcriptase polymerase chain reaction and expression arrays. Now single nucleotide polymorphisms, array comparative genomic hybridization, methylation signatures, microRNAs, and proteomics are increasingly included in biomarker profiling; nanotechnology, molecular imaging, and the use of circulating tumor cells are advancing at a tremendous pace. By the time this book is released, the specimen collection guidelines for breast cancer, representing a joint effort between the North American cooperative
groups and those comprising the Breast International Group (BIG), will have been published and be available for open Web access. We have attempted to cover the breadth of these diverse areas in this book.

I am especially proud of three aspects of the book. The first is the timeliness, as so many new approaches to individualization of therapy have emerged, and when two very large randomized trials are aggressively accruing on the basis of Oncotype and mammoprint. Second, the diversity of the topics covered, ranging from pathology, tissue handling, DNA-, expression-, and epigenetic-based approaches, nanotechnology, stem cells, circulating tumor cells, endocrine, and chemotherapeutic through to pharmacologic aspects and the key pivotal trials. Third, I am proud that so many of my most distinguished friends and colleagues from around the world have been so generous with their time and devotion to make this book outstanding.

Finally, in writing this preface from one of the most beautiful places on earth, I am deeply reminded of the disparities in access to healthcare. We are all immensely grateful that my friends and colleagues John Seffrin and Otis Brawley at the American Cancer Society have been leaders in bringing this critical issue to the forefront of public consciousness, but we must never fail to remember that severe disparities in access to care remain, not only in the United States but globally. Individualized therapy is not only meant to direct therapy to those who will benefit most but also intended to reduce waste and improve cure rates dramatically, so that this therapy is available to all who suffer. Let us, the privileged, never move our focus from the most disadvantaged in our world and aim at using individualized therapy as a means to make our most sophisticated treatment available and affordable to all.

Brian Leyland-Jones, M.D., Ph.D.
Acknowledgments

Many thanks are owed on a personal level, especially to all my friends and coauthors who have been immensely generous in dedicating their nonexistent spare moments to contributing the various chapters. I am deeply grateful to the publisher, Informa Healthcare, who not only had the insight to partner with me on this book but also had the patience to wait for the “nonexistent spare moments” of the contributing authors to materialize. I must also profoundly apologize to several colleagues who were approached early on to contribute chapters but who were unable to because of the original time lines and now must be frustrated because the time lines were forced to be extended considerably. Most especially, I must thank my dear, dear colleague, Dr. Brian R. Smith, who was relentless in organization, follow-up, and advice, a man whose wisdom I rely upon at so many critical moments and to whose structured approach I aspire. I am deeply, deeply grateful to Charles Bronfman, whose chair in the memory of his dear sister, Minda, sustained me over the past 17 years and who now has inspirationally established his own Institute of Personalized Cancer Therapy in New York. I am grateful also to the Flanders family who have shown me a depth of kindness, support, and friendship that is without equal. Finally, I am humbled by my research and clinical colleagues who instruct and guide me constantly, and especially by each and every patient who inspires us and teaches, in each and every encounter, just how personalized breast cancer therapy really is.
Contents

Preface . . . . . . . . . . iii
Acknowledgments . . . . v
Contributors . . . . xi

1. Pharmacogenetics of Breast Cancer: Toward the Individualization of Therapy .......................... 1
   Jenny C. Chang, Susan G. Hilsenbeck, and Suzanne A. W. Fuqua

2. Pharmacology, Pharmacogenetics, and Pharmacoepidemiology: Three P’s of Individualized Therapy ............... 11
   Shaheenah Dawood

3. Role of Genetic Variability in Breast Cancer Treatment Outcomes ........................................... 25
   Kandace L. Amend, Ji-Yeob Choi, and Christine B. Ambrosone

4. The Role of Epigenetics in Breast Cancer: Implications for Diagnosis, Prognosis, and Treatment ..................... 45
   Amy M. Dworkin, Tim H.-M. Huang, and Amanda E. Toland

5. Comparative Genomic Hybridization and Copy Number Abnormalities in Breast Cancer .......................... 61
   Nicholas Wang and Joe Gray
6. Gene Expression Profiling as an Emerging Diagnostic Tool to Personalize Chemotherapy Selection for Early Stage Breast Cancer .......................... 77
   Cornelia Liedtke and Lajos Pusztai

7. Gene Expression–Based Predictors of Prognosis and Response to Chemotherapy in Breast Cancer .................................................. 97
   Soonmyung Paik

8. Utilization of Genomic Signatures for Personalized Treatment of Breast Cancer ................................................................. 107
   P. Kelly Marcom, Carey K. Anders, Geoffrey S. Ginsburg,
   Joseph R. Nevins, and Anil Potti

9. Personalized Medicine by the Use of Microarray Gene Expression Profiling ................................................................. 123
   Stella Mook, Laura J. van ’t Veer, and Fatima Cardoso

10. Predictive Markers for Targeted Breast Cancer Treatment ......... 135
    Hans Christian B. Pedersen and John M. S. Bartlett

11. Circulating Tumor Cells in Individualizing Breast Cancer Therapy ................................................................. 151
    James M. Reuben and Massimo Cristofanilli

12. Individualization of Endocrine Therapy in Breast Cancer ....... 167
    Amelia B. Zelnak, Ruth M. O’Regan, and Clodia Osipo

13. TAILORx: Rationale for the Study Design .............................. 185
    Joseph A. Sparano

14. Taking Prognostic Signatures to the Clinic ............................ 197
    Michail Ignatiadis and Christos Sotiriou

15. Tissue is the Issue ................................................................. 213
    Adekunle Raji

16. Acquisition and Preservation of Tissue for Microarray Analysis ................................................................. 231
    Brian R. Untch and John A. Olson
Contents ix

17. Tapping into the Formalin-Fixed Paraffin-Embedded (FFPE)
   Tissue Gold Mine for Individualization of Breast
   Cancer Treatment ................................................. 243
   Mark Abramovitz and Brian Leyland-Jones

18. Breast Cancer Stem Cells and Their Niche: Lethal Seeds in
   Lethal Soil .......................................................... 259
   Danuta Balicki, Brian Leyland-Jones, and Max S. Wicha

19. Molecular Imaging in Individualized Cancer Management ... 291
   David M. Schuster and Diego R. Martin

20. Cancer Nanotechnology for Molecular Profiling and
   Individualized Therapy ........................................... 309
   May Dongmei Wang, Jonathan W. Simons, and Shuming Nie

Index .......... 323
Contributors

Mark Abramovitz  VM Institute of Research, Montreal, Québec, Canada

Christine B. Ambrosone  Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, New York, U.S.A.

Kandace L. Amend  Department of Oncological Sciences, Mount Sinai School of Medicine, New York, New York, U.S.A.

Carey K. Anders  Duke Institute for Genome Sciences & Policy, Department of Medicine, Duke University Medical Center, Durham, North Carolina, U.S.A.

Danuta Balicki  Department of Medicine and Research Centre, Centre Hospitalier de l’Université de Montréal; Department of Medicine, Université de Montréal; and Breast Cancer Prevention and Genetics and Breast Cancer Risk Assessment Clinic, Ville Marie Multidisciplinary Breast Cancer Centre, Montréal, Québec, Canada

John M. S. Bartlett  Endocrine Cancer Group, Cancer Research Centre, Western General Hospital, Edinburgh, United Kingdom

Fatima Cardoso  Department of Medical Oncology, Institute Jules Bordet, Brussels, Belgium

Jenny C. Chang  Breast Center, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, U.S.A.

Ji-Yeob Choi  Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, New York, U.S.A.
Contributors

Massimo Cristofanilli  Department of Breast Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, U.S.A.

Shaheenah Dawood  Department of Medical Oncology, Dubai Hospital, Dubai, United Arab Emirates

Amy M. Dworkin  Human Cancer Genetics Program, Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University, Columbus, Ohio, U.S.A.

Suzanne A. W. Fuqua  Breast Center, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, U.S.A.

Geoffrey S. Ginsburg  Duke Institute for Genome Sciences & Policy, Department of Medicine, Duke University Medical Center, Durham, North Carolina, U.S.A.

Joe Gray  Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.

Susan G. Hilsenbeck  Breast Center, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, U.S.A.

Tim H.-M. Huang  Human Cancer Genetics Program, Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University, Columbus, Ohio, U.S.A.

Michail Ignatiadis  Department of Medical Oncology, Translational Research Unit, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium

Brian Leyland-Jones  Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia, U.S.A.

Cornelia Liedtke  Department of Breast Medical Oncology, The University of Texas, Anderson Cancer Center, Houston, Texas, U.S.A.

P. Kelly Marcom  Duke Institute for Genome Sciences & Policy, Department of Medicine, Duke University Medical Center, Durham, North Carolina, U.S.A.

Diego R. Martin  Department of Radiology, Emory University, Atlanta, Georgia, U.S.A.
Contributors

Stella Mook  Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Joseph R. Nevins  Duke Institute for Genome Sciences & Policy, Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina, U.S.A.

Shuming Nie  Departments of Biomedical Engineering and Chemistry and the Winship Cancer Institute, Emory University and Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

Ruth M. O’Regan  Emory Winship Cancer Institute, Atlanta, Georgia, U.S.A.

John A. Olson  Institute for Genome Sciences and Policy and the Department of Surgery, Duke University Medical Center, Durham, North Carolina, U.S.A.

Clodia Osipo  Loyola University, Maywood, Illinois, U.S.A.

Soonmyung Paik  Division of Pathology, NSABP Foundation, Pittsburgh, Pennsylvania, U.S.A.

Hans Christian B. Pedersen  Endocrine Cancer Group, Cancer Research Centre, Western General Hospital, Edinburgh, United Kingdom

Anil Potti  Duke Institute for Genome Sciences & Policy, Department of Medicine, Duke University Medical Center, Durham, North Carolina, U.S.A.

Lajos Pusztai  Department of Breast Medical Oncology, The University of Texas, Anderson Cancer Center, Houston, Texas, U.S.A.

Adekunle Raji  Eastern Cooperative Oncology Group Pathology Coordinating Office, Northwestern University, Evanston, Illinois, U.S.A.

James M. Reuben  Department of Hematopathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, U.S.A.

David M. Schuster  Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, Emory University, Atlanta, Georgia, U.S.A.

Jonathan W. Simons  The Prostate Cancer Foundation, Santa Monica, California, U.S.A.
Contributors

Christos Sotiriou  Department of Medical Oncology, Translational Research Unit, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium

Joseph A. Sparano  Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, New York, U.S.A.

Amanda E. Toland  Human Cancer Genetics Program, Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University, Columbus, Ohio, U.S.A.

Brian R. Untch  Institute for Genome Sciences and Policy and the Department of Surgery, Duke University Medical Center, Durham, North Carolina, U.S.A.

Laura J. van ’t Veer  Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

May Dongmei Wang  Departments of Biomedical Engineering and Electrical and Computer Engineering, Georgia Institute of Technology and Emory University, Atlanta, Georgia, U.S.A.

Nicholas Wang  Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.

Max S. Wicha  Comprehensive Cancer Centre, Department of Internal Medicine, Division of Hematology-Oncology, University of Michigan, Ann Arbor, Michigan, U.S.A.

Amelia B. Zelnak  Emory Winship Cancer Institute, Atlanta, Georgia, U.S.A.
Pharmacogenetics of Breast Cancer: Toward the Individualization of Therapy

Jenny C. Chang, Susan G. Hilsenbeck, and Suzanne A. W. Fuqua

Breast Center, Dan L. Duncan Cancer Center,
Baylor College of Medicine, Houston, Texas, U.S.A.

INTRODUCTION

Mortality from breast cancer results from the ability of some tumors to metastasize to distant sites. Selecting patients with micrometastases at diagnosis is crucial for clinicians in deciding who should, and who should not, receive toxic and expensive adjuvant chemotherapy to eradicate these metastatic cells. Although many individual biomarkers were originally attractive, over the years most have failed to become clinically useful. In addition, the management of breast cancer has changed, with the majority of node-negative patients now undergoing systemic adjuvant therapy because we cannot precisely determine an individual’s risk of recurrence. A majority of node-negative patients are being unnecessarily overtreated because if left systemically untreated, only about 25% of node-negative patients would ever develop recurrence. There is therefore a critical need to identify patients with sufficiently low risk of breast cancer recurrence so as to avoid further treatment. In addition, in patients at risk of recurrence and in need of therapy, optimal therapeutic selection is an increasingly important objective. Recent developments in applying microarray technologies to breast tumor samples suggest that these new techniques may provide for the transition of molecular biological discoveries to clinical application and...
will generate clinically useful genomic profiles that more accurately predict the long-term outcome of individual breast cancer patients.

**BACKGROUND**

Until recently, evaluations of prognostic and predictive factors have considered one factor at a time or have used small panels of markers. However, with the advent of new genomic technologies such as microarrays, capable of simultaneously measuring thousands of genes or gene products, we are beginning to construct molecular fingerprints of individual tumors so that accurate prognostic and predictive assessments of each cancer may be made. Clinicians may one day base clinical management on each woman’s personal prognosis and predict the best individual therapies according to the genetic fingerprint of each individual cancer.

Breast cancer is characterized by a very heterogeneous clinical course. A major goal of recent studies is to determine whether RNA microarray expression profiling or DNA array gene amplification or gene loss patterns can accurately predict an individual’s long-term potential for recurrence of breast cancer, so that appropriate treatment decisions can be made. Microarrays can be used to measure the mRNA expression of thousands of genes at one time or survey genomic alterations that may distinguish molecular phenotypes associated with long-term, recurrence-free survival or clinical response to treatment. These new technologies have been successfully applied to primary breast cancers and may eventually outperform currently used clinical parameters in predicting disease outcome.

Since RNA expression microarray technology provided a method for monitoring the RNA expression of many thousands of human genes at a time, there was considerable anticipation that it would quickly and easily revolutionize our approaches to cancer diagnosis, prognosis, and treatment. The reality remains extremely promising but is also complex. A potential complication in the application of microarray technology to primary human breast tumor samples is the presence of variable numbers of normal cells, such as stroma, blood vessels, and lymphocytes, in the tumor. Indeed, it has been demonstrated using gross analysis of human breast cancer specimens compared with breast cancer cell lines that the tumors expressed sets of genes in common not only with these cell lines but also with cells of hematopoietic lineage and stromal origin (1,2). Laser capture microdissection has also been successfully used to isolate pure-cell populations from primary breast cancers for array profiling (3). Sgroi et al. (3) utilized laser capture microdissection to isolate morphologically “normal” breast epithelial cells, invasive breast cancer cells, and metastatic lymph node cancer cells from one patient and were able to demonstrate the feasibility of using microdissected samples for array profiling as well as following potential progression of cancer in this patient. However, with the emerging data supporting important roles for the surrounding stroma in breast cancer progression and the
Pharmacogenetics of Breast Cancer

labor-intensive and technically challenging nature of laser capture technology with subsequent amplification of RNA for quantitation, most published investigations to date have evaluated total gene expression to identify prognostic profiles, as will be described in the next section.

MOLECULAR CLASSIFICATION OF BREAST CANCER

A study of sporadic breast tumor samples by Perou et al. (2) was the first to show that breast tumors could be classified into subtypes distinguished by differences in their expression profiles. Using 40 breast tumors and 20 matched pairs of samples before and after doxorubicin treatment, an “intrinsic gene set” of 476 genes was selected that was more variably expressed between the 40 sporadic tumors than between the paired samples. This intrinsic gene set was then used to cluster and segregate the tumors into four major subgroups: a “luminal cell-like” group expressing the estrogen receptor (ER); a “basal cell-like” group expressing keratins 5 and 17, integrin β4, and laminin, but lacking ER expression; an “Erb-B2-positive” group; and a “normal” epithelial group (Fig. 1).

In a subsequent study with 38 additional cancers, the investigators found the same subgroups as before (4), except that the luminal, ER-positive group was further subdivided into subsets with distinctive gene expression profiles. In univariate survival analysis, performed on the 49 patients diagnosed with locally advanced disease but without evidence of distant metastasis, ER positivity was not a significant prognostic factor on its own, but the luminal-type group enjoyed a more favorable survival compared with the other groups. Conversely, the basal-like group had a significantly poorer prognosis. Although small and exploratory, this study suggests that important differences in outcome can be ascertained from microarray expression profiling.

An interesting study was reported by Gruvberger et al. (5), who profiled 58 grossly dissected primary invasive breast tumors and used artificial neural network analysis to predict the ER status of the tumors on the basis of their gene expression patterns. They then determined which specific genes were the most important for ER classification. By comparing to Serial Analysis of Gene Expression (SAGE) data from estradiol-stimulated breast cancer cells, they determined that only a few genes of the many genes that were associated with ER expression in tumors were indeed estrogen responsive in cell culture. This observation lent further support to the hypothesis developed by Perou et al. (1) that basic cell lineages, such as the luminal ER-positive cell type, can be partly explained by observed genomic gene expression patterns rather than downstream effectors of only one pathway, such as the ER.

PROGNOSTIC IMPLICATIONS

Microarrays have also been used to predict lymph node status and very short-term relapse-free survival in two groups (n = 37 and 52, respectively) of heterogeneously treated patients (6). Although prediction of nodal status is of
limited interest clinically, the study uses innovative statistical methods, rigorously generates estimates of future classifier performance, and further demonstrates the feasibility of accurate prediction of tumor biology using expression arrays. In a more focused and somewhat more clinically relevant study, van ‘t Veer et al. (7) used RNA expression microarray analyses to identify a 70-gene prognostic gene signature (“classifier”) in young, untreated, axillary lymph node-negative patients using a training set of 44 good (disease free for more than 5 years) and 34 poor (distant relapse in less than 5 years) outcome tumors and then tested the classifier in a validation set of 19 tumors. The same group (8) has now extended the study to a total of 295 young (<53 years), stage I–II breast cancer patients with both node-negative and node-positive disease, using the

Figure 1  Supervised classification of prognosis signatures (van’t Veer). The 78 tumors are listed vertically, and the 70 “prognostic” genes horizontally. (A) Study design. (B) Heat map of prognostic signature.
Pharmacogenetics of Breast Cancer

70-gene classifier (7). The microarray-based predictions are consistent with, and perhaps better than, estimates that can be obtained with current prognostic indices.

GENETIC SUSCEPTIBILITY

A few studies have utilized new genomic approaches for the study of inherited breast cancer (8). There is accumulating evidence, both epidemiological and histological, that tumors arising as a result of mutations in the two breast cancer susceptibility gene families (BRCA1 and BRCA2) are biologically distinct. For instance, BRCA1 breast cancers are most often ER and PR negative, but BRCA2 cancers more often tend to be positive for these receptors (9). In a seminal paper published by Hedenfalk et al. (10), seven tumors each from BRCA1 and BRCA2 gene mutation carriers or sporadic breast cancers were compared by expression microarray analysis. They found that the gene expression profiles of the three tumor groups differed significantly from each other, underscoring the fundamental differences between BRCA1 and BRCA2 mutation-associated tumors. Of course a potential confounding issue was the differential distribution of ER between the BRCA1 and BRCA2 tumors. However, even after removal of ER- or PR-associated genes from the analysis, the two inherited tumor groups were still discernable. Thus, ER status alone does not fully explain the observed differences in gene expression profiles. Although this study is obviously very small, and other confounding issues such as tumor stage, grade, and treatment cannot be considered, it does set a foundation for larger validation studies to confirm differential genes, which could then provide important clues to the etiology of inheritable breast cancer.

Microarrays are also being studied as a way to predict response to systemic therapy. The neoadjuvant setting is especially attractive for these studies for several reasons, including early assessment of response to therapy, biopsiable access to the primary tumor, and considerable reduced sample sizes compared to those required in the adjuvant setting.

PREDICTIVE IMPLICATIONS

Methods for assessing response in neoadjuvant trials remain problematic. Clinical response to neoadjuvant chemotherapy is a validated surrogate marker for improved survival (11,12). Women who achieve pathologic complete response are most likely to have the best clinical outcome, although survival is still improved in those who clinically respond but do not achieve pathologic complete response.

In an early study, Buchholz et al. (13) obtained sufficient RNA from core biopsies of five patients to perform serial microarray expression profiles and showed that despite differences in therapy, patients with good pathological
responses to neoadjuvant treatment appeared to have gene profiles that clustered distinctly than those of patients who were poor responders to treatment. More recently, Chang et al. (14) have shown that gene profiling can be used to accurately predict response to neoadjuvant docetaxel. The study enrolled 24 subjects, extracted sufficient RNA from all core needle biopsies, and constructed a 92-gene predictor of response. In a complete cross-validation analysis, which gives an unbiased estimate of performance on future samples, the classifier correctly identified 10 of 11 responders and 11 of 13 nonresponders with an overall accuracy of 88%. In a small validation set, this 92-gene classifier successfully predicted response in six patients (Fig. 2). This compares very favorably with the best existing predictive factors for response to specific therapy and strongly suggests that after appropriately extensive validation, microarray profiling will be useful for treatment selection. A second neoadjuvant study was recently published that used cDNA arrays to develop predictors for paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide, involving 24 samples. A classifier with 74 markers was developed, with 78% accuracy, suggesting that transcriptional profiling has the potential to identify a gene expression pattern in breast cancer that may lead to clinically useful predictors of chemotherapy response (15).

In vitro drug sensitivity data in NCI60 cell lines sensitive and resistant to specific chemotherapeutic agents (adriamycin, cyclophosphamide, docetaxel, etoposide, 5-fluorouracil, paclitaxel, topotecan) were recently published (16). Gene set enrichment analysis was performed, and gene expression signatures predictive of sensitivity to individual chemotherapeutic drugs were reported. Each signature was validated with response data from an independent set of cell line studies. These signatures were then used to predict clinical response in individuals treated with these drugs. Notably, signatures developed to predict response to individual agents, when combined, were found to also predict response to multidrug regimens (16). These profiles are being validated in prospective clinical trials.

However, we acknowledge that studies to construct and validate array-based prognostic and predictive “markers” are complex. These studies must address all of the concerns associated with ordinary, single-gene markers as well as a number of considerations unique to array studies. Recommendations for development of array-based prognostic classifiers have recently been enunciated by Simon et al. (17). Among the most important points, they recommend that studies should include (1) adequately large sample sizes in both training and validation sets, (2) complete iteration of the entire classifier construction process in estimating cross-validated prediction rates, (3) head-to-head comparison of alternative classifiers on the same dataset, and (4) the full diversity of cases in any validation set. In addition, gene expression patterns may be confounded by several other factors, including ovarian ablation in premenopausal, ER-positive patients and a different mechanism of action of combination chemotherapies.
Groups of patients with different characteristics, such as menopausal and ER status or human epidermal growth factor receptor 2 (HER2) overexpression, may be necessary to definitively determine classifying patterns in these subsets of patients.

Figure 2  Hierarchical clustering of genes correlated with docetaxel response. Sensitive tumors (S) are defined as 25% residual disease, and resistant tumors (R) are defined as greater than 25% residual disease.
CONCLUSIONS

The goal of comprehensive, genomewide approaches is to identify clinically useful genetic profiles that will accurately predict the outcome to therapy and prognosis of patients with breast cancer. Despite improvements in technology, complex mechanisms driving breast cancer evolution continue to present challenges for the use of genomic approaches to better understand breast and other cancers. This, combined with our use of different markers, methods, tumors (e.g., differing ER and HER2), and measurements of clinical outcomes, impedes the development of a consensus regarding predictive and prognostic markers for breast cancer.

As this field matures, genomic studies examining identical breast tumor sets with multiple complementary technologies (e.g., loss of heterozygosity, comparative genomic hybridization, gene expression array analysis) will prove essential to unraveling the genetic heterogeneity characteristic of this disease. A combined genomic approach is necessary to define the underlying heterogeneous complexity characteristic of breast cancer. These data should lead to the identification and characterization of breast cancer subtypes, define the malignant potential of a given lesion, and predict its sensitivity to specific therapies. These multidisciplinary approaches should contribute to a better biological understanding, and therefore improved clinical management, of breast cancer.

ACKNOWLEDGMENTS

This work was supported by funding provided by National Institutes of Health (NIH) grants P50 CA58183 and P01 CA30195, the Breast Cancer Research Foundation, the Emma Jacobs Clinical Breast Cancer Fund, and the U.S. Army Medical Research and Materiel Command DAMD17-01-0132 and W81XWH-04-1-0468.

REFERENCES

Pharmacogenetics of Breast Cancer

Pharmacology, Pharmacogenetics, and Pharmacoepidemiology: Three P’s of Individualized Therapy

Shaheenah Dawood
Department of Medical Oncology, Dubai Hospital, Dubai, United Arab Emirates

If it were not for the great variability among individuals medicine might as well be a science and not an art.

Sir William Osler, 1892

INTRODUCTION

Over the last decade significant advances have been made in the management of breast cancer. This is exemplified by the development of several new agents and significantly improved survival rates. However, despite our best efforts breast cancer remains the second most common cause of cancer-related death in the United States with approximately 40,910 women expected to die of this disease in 2007 (1). Unfortunately, such trends are not restricted to only breast cancer but have been observed across a range of solid tumors. A large portion of the underlying problem is attributed to the fact that the choice of a chemotherapeutic agent is often empirical based on guidelines that assumes all patients with a particular malignancy to be homogeneous. The unfortunate outcome of this
erroneous belief is that approximately 40% of patients may be receiving the wrong drug (2).

Numerous studies have consistently observed that the efficacy and toxicity of chemotherapeutic agents are heterogeneous rather than homogeneous across populations. Despite a multitiered approach to determining the efficacy and safety profiles of chemotherapeutic agents, spanning from in vitro assays, animal studies through to large human population studies, the current practice of administering the same dose of a given anticancer drug to a population of patients results in considerable variation in clinical outcome. These effects cover a spectrum of possibilities ranging from success to failure when considering treatment outcome as an endpoint, and from no effect to a lethal event when treatment-associated toxicity is the variable under study (3). Complicating the problem further is the fact that currently available anticancer drugs show limited efficacy in up to 70% of patients with adverse drug reactions accounting for as many as 100,000 patient deaths and two million hospitalizations in the United States every year (4).

The key is to be able to, at the individual level, prospectively identify patients at risk for severe toxicity and those likely to attain maximum benefit from a given chemotherapeutic agent. Understanding that an individual does not live in isolation but is in fact a product of the gene-environment interaction is also vital to our understanding of the concept of individualized therapy (Fig. 1).

Figure 1 Sources of pharmacologic and pharmacogenetic variability.
Designing treatment strategies based on individual efficacy and toxicity predictions will shift clinical practice from simply an “art” to a “science” by incorporating a combination of physiological variables, genetic characteristics, and environmental factors (all known to alter the dose-concentration time profile) as well as population epidemiological elements resulting in a practice paradigm of “personalized medicine.”

THE COMPONENTS OF INDIVIDUALIZED THERAPY: THE THREE P’s

When we administer a drug to a patient, our goal is to deliver a dose that will achieve maximum benefit with little, if any, associated toxicity. Seldom is this goal achieved. Individual differences in tumor response and normal tissue toxicity have been consistently observed with most anticancer agents. These differences are due to a variety of clinical variables (e.g., gender, age, diet, and drug-drug interactions) as well as variations in drug metabolism and transport. It is believed that 30% of patients treated with a particular drug will exhibit beneficial effects from the drug, 30% will not attain any form of beneficial effect, 10% will experience only side effects from the drug, and 30% are non-compliant with the drug often as a result of drug-related side effects or perceived insufficient effect (5). This results in a waste of allocated resources with large proportions of patients being unnecessarily exposed to adverse side effects. Here we will describe three components that aim to explain the observed differences in responses to individual drugs: pharmacology, pharmacoepidemiology, and pharmacogenetics.

Current dosing of anticancer drugs is based on either a fixed quantity of the drug or on a dose that is normalized to the individual body surface area. This method makes the assumption that within a group of individuals there will be a uniform degree of systemic exposure to the drug. With studies reporting 2- to 10-fold variations in drug clearance, this assumption is clearly not valid (6). With drugs used in the field of oncology having narrow therapeutic indices, it thus becomes imperative that we understand the mechanisms behind the observed variability in drug response and toxicity when treating a patient with cancer. At the clinical level, variability in drug response can be explained by its pharmacology that describes the pharmacokinetic and pharmacodynamic profiles of the drug. Pharmacokinetics explain “what the body does to the drug” by describing the relationship between time and plasma concentrations of the drug metabolites. This relationship is affected by variables such as absorption, distribution, metabolism, and excretion of the drug. The underlying objective of pharmacodynamics is to describe “what the drug does to the body” by correlating drug concentration to drug effect (both beneficial and adverse effects). Both the pharmacokinetic and pharmacodynamic components of a drug are not independent, but represent a spectrum of continuous events starting with the ingestion of the drug and culminating in the observed clinical effect.
Variability in any of the pharmacokinetic components will affect the observed drug effect. This is typically observed in patients with organ dysfunction where the inability to either metabolize or excrete the drug will lead to exaggerated drug effects. For example, the use of carboplatin in patients with renal dysfunction exaggerates the thrombocytopenic effect of this drug. Dosing formulas using creatinine clearance are used to avoid this problem while maximizing dose intensity (7). Another example is the use of taxanes, which are primarily metabolized and excreted by the liver. When administered to patients with hepatic dysfunction, prolonged bouts of severe neutropenia are observed. These examples illustrate how clinical variables can affect the pharmacokinetic profile of a drug leading to disturbances in effect. Good clinical assessment of the patients is therefore the first and most important step in designing an individual treatment regimen that will maximize drug benefit.

Pharmacoepidemiology is a sub discipline of epidemiology that focuses on understanding why individuals respond differently to drug therapy. It is defined “as the study of the distribution and determinants of drug-related events in populations and the application of this study to efficacious drug treatments” (8). These drug-related events pertain to both beneficial and adverse events. Pharmacoepidemiology, through the use of observational methods, aims to quantify drug exposure and to describe, as well as predict, the effectiveness and safety of the drug within defined populations at specific places and time. Factors important in interpreting the variability in drug outcome include a patient’s health profile, sex, age, diet, associated comorbidities, disease severity and prognosis, drug compliance, drug prescribing and dispensing quality, and genetic profile of the patient and tumor (9,10). Depending on drug effect there are four main groups of patients: (i) responders, (ii) non-responders, (iii) toxic responders, and (iv) toxic non-responders (Fig. 2).

Pharmacogenetics, a third and important component of individualized therapy, focuses on describing the extent to which an individual’s genetic makeup is responsible for the observed differences in drug efficacy and toxicity profiles (11,12). This information is then used to make predictions about the safety, toxicity, and efficacy of drugs in individual patients. Inherited variability of drug targets, drug metabolizing enzymes, and drug transporters may all have a major impact on overall drug response, disposition, and associated adverse side effects by altering the pharmacokinetic and pharmacodynamic properties of drugs (Table 1).

Genetic variations include nucleotide repeats, insertions, deletions, and single nucleotide polymorphisms (SNPs). SNPs are the simplest and most commonly studied DNA polymorphism that occur once every 1900 base pairs in the 3 billion bases in the human genome (13) and account for more than 90% of the genetic variation observed. More than 1.4 million SNPs have been identified in the human genome. Of these more than 60,000 SNPs occur in the coding region of genes with some SNPs being associated with variations in drug metabolism and effects (13). When polymorphisms occur in the promoter region
it can affect the transcription of the gene. When the polymorphism occurs in the exon (and intron) region or 3'UTR of the gene, it can affect translation and RNA stability, respectively, resulting in either reduced or enhanced activity of the encoded protein (3).

Two approaches have been used to identify clinically relevant polymorphisms. In the reverse genetics approach, a drug is first given to a group of individuals and the variability in response to and disposition of the drug is observed. Phenotypes of clinical importance are prioritized and then experiments
are conducted to determine if a genetic component is present to explain the observed variability. Finally, the biochemical and molecular nature of the variability is determined through extensive resequencing techniques that will link the polymorphisms with the observed phenotypes. The second and more cost effective method is the forward genetics approach. This method involves first determining the presence of genetic variability at specific loci in a group of individuals. A hypothesis is then formulated regarding these polymorphisms causing phenotypic variability. Finally, an experiment is performed to investigate the clinical consequences of the genetic variability. Because of the multitude of polymorphisms, the challenge today is to be able to identify polymorphisms that occur in gene regulatory or coding regions that have clinical relevance. The field of oncology adds another level of complexity as issues such as clinically relevant tumor genome variations not found in the germline and somatic mutations acquired before or after chemotherapy can affect drug efficacy and toxicity. Moreover, different polymorphisms of certain genes may exhibit different phenotypes among different malignancies. Polymorphisms of glutathione S-transferases (GSTs) are a common example. GSTs facilitate the conjugation of glutathione to xenobiotics, including anticancer agents (e.g., cyclophosphamide and anthracyclines), environmental carcinogens, and reactive oxygen products (14). A variety of functionally important polymorphisms of GSTs have been described. The presence of the GSTT1 polymorphism has been shown to predispose children treated intensively for acute myelogenous leukemia to greater toxicity and death during remission (15), while women with breast cancer treated with cyclophosphamide carrying the GSTP1 polymorphism are more likely to survive (16). This example illustrates how the phenotypic effect of polymorphisms depends on the underlying tumor as well as the intensity of regimen used to treat the malignancy. However, as we begin to unravel and accurately identify the mechanism of action of many anticancer drugs, polymorphisms in candidate genes likely to influence drug efficacy and associated adverse events can be identified. Indeed the field of pharmacogenetics has transitioned from one that determined the genetic basis of observed inter patient phenotypic variability to one that determines the phenotypic consequence of an observed genotypic variability.

Linking the information obtained from the pharmacology and pharmacoepidemiology of a drug and the pharmacogenetics that describes the genetic makeup of the individual being treated with the drug is thus vitally important in developing an optimal approach to individualized therapy. The use of isoniazid in the treatment of tuberculosis is a good example. Epidemiological evidence shows that polymorphisms differ in frequency among ethnic and racial groups. This phenomenon was shown early on when interindividual and interethnic variations of efficacy and toxicity associated with isoniazid treatment was observed in a cohort of individuals with tuberculosis. The enzyme responsible for the metabolism of isoniazid is N-acetyltransferase, a phase II conjugating liver enzyme. Early studies of isoniazid revealed a bimodal distribution in its plasma concentration among individuals. High plasma concentrations were
observed among individuals with the slow acetylator phenotype often resulting in peripheral nerve damage, while low plasma concentrations of isoniazid were observed among individuals with fast acetylator phenotype and were not affected (17). Studies have shown large variations in the incidence of the slow acetylator phenotype (inherited as an autosomal recessive trait) among different ethnic groups (40–70% of Caucasians and African Americans, 10–20% of Japanese, and more than 80% of Egyptians) (18,19). Amonafide (20), a topoisomerase II inhibitor, is another substrate of N-acetyltransferase whose clinical development has been hampered by highly variable and unpredictable toxicity caused, at least, in part, by interindividual differences in N-acetylation with fast acetylators experiencing greater myelosuppression than slow acetylators. This example illustrates both the importance of pharmacoepidemiological evidence in identifying and characterizing phenotypic variance to individual drugs and how the pharmacological effects of these drugs depend on pharmacogenetic determinants that define sensitivity to the drug, and ultimately describes the drug’s behavior.

THE LEUKEMIA STORY: WHAT DID WE LEARN?

Pharmacological and pharmacogenetic studies are increasingly being used to optimize existing drug therapy with the goal of reducing toxicity while maximizing efficacy, leading to improved survival. By utilizing this method remarkable progress has been observed in the treatment of acute lymphoblastic leukemia (ALL) with cure rates increasing from a mere 10% (four decades ago) to nearly 80% in children and 40% in adults (21). This has been achieved through optimization of use of a variety of commonly used antileukemic drugs such as mercaptopurine and methotrexate.

Thiopurine methyltransferase (TPMT) is an enzyme that converts thiopurine prodrugs, such as mercaptopurine, into inactive methylated metabolites. Prior to inactivation, mercaptopurine is converted into a variety of active nucleotide metabolites that are incorporated into DNA and RNA, thereby inducing an antileukemic effect. Patients with an inherited TPMT deficiency suffer severe, potentially fatal hematopoietic toxicity when exposed to standard doses of mercaptopurine (and other thiopurine prodrugs such as azathioprine and thioguanine). More than 95% of the clinically relevant TPMT mutations are accounted for by three nonsynonymous SNPs (TPMT*2, TPMT*3A, and TPMT*3C) (22). TPMT activity is inherited as an autosomal codominant trait with approximately 95% of the population homozygous for the wild-type allele TPMT*1 (that have full enzyme activity), 10% heterozygous for the polymorphism (that have intermediate levels of enzyme activity), and 1 in 300 individuals will carry two mutant TPMT alleles (that do not express functional TPMT) (23). Using a pharmacogenetic test, developed at St. Jude Hospital, patients are often classified according to their levels of TPMT activity into normal, intermediate, and deficient. This test has a concordance rate of 100% between genotype and phenotype with studies revealing no long-term outcome
changes related to dosage individualization based on the TPMT status of a patient (24) (Fig. 3). This illustrates how the clinical application of pharmacogenetics can be used to minimize toxicity without compromising efficacy.

Methyltetrahydrofolate reductase (MTHFR), a crucial enzyme in the metabolism of folic acid, catalyses the reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate (the predominant circulatory form of folate and the carbon donor for the demethylation of homocysteine to methionine). As methotrexate increases serum levels of homocysteine and the active polyglutamate metabolites of methotrexate inhibit MTHFR function, patients with a low level of MTHFR function are at an increased risk of methotrexate-induced toxicities such as oral mucosities. Two polymorphisms associated with reduced MTHFR activity have been described (25): (i) 677C > T polymorphism that is homozygous in 10% of the population where it encodes the enzyme with 30% of the wild-type enzyme activity and heterozygous in 40% of the population where it encodes an enzyme that has 60% of the wild type-enzyme activity and (ii) 1298A > C (less common than 677C > T). Recent findings have also suggested that MTHFR polymorphisms may protect against the development of both adult and pediatric ALL (26).

5-FLUOROURACIL

5-Fuorouracil (5-FU) is a uracil analog widely used to treat solid tumors such as breast and colorectal cancer (27). 5-FU is inactivated by dihydropyrimidine dehydrogenase (DPD), an enzyme that exhibits up to 20-fold variation in activity among individuals (28). Mutations that result in low DPD activity predispose to potentially life-threatening gastrointestinal, hematopoietic, and neurological toxicities (29). Up to 20 mutations (30) have been identified that cause inactivation of DPD. In the general population it has been estimated that 3% to 5% of individuals are heterozygous and 0.1% are homozygous carriers for the mutation that inactivates DPD, with DPYD*2A being the most inactivating allele (31). Although polymorphisms associated with DPD inactivation have been identified, a number of patients still exhibit severe 5-FU toxicity with no detectable
mutations in the coding region of the DPD gene explaining the difficulty in implementation of DPD pharmacogenetics in the prospective identification of high-risk patients for severe 5-FU toxicity.

One of the primary mechanisms of action of 5-FU is inhibition, via fluorodeoxyuridine monophosphate (an active metabolite of 5-FU), of thymidylate synthase (TS), a critical enzyme required for de novo thymidylate synthesis, which is required for DNA synthesis and repair (32). Polymorphisms in the gene encoding TS are one of the few examples of a polymorphism that is associated with the genetic makeup of the tumor with clinical studies showing an inverse relationship between TS levels and antitumor response (33). The TSER polymorphism of TS is characterized by tandem repeats in the TS promoter enhancer region of the gene. Higher levels of TS expression and enzyme activity have been observed with increasing copies of tandem repeats, which result in lower tumor response rates to 5-FU (34). In a study by Marsh et al (35), it was observed that in a cohort of patients with metastatic colorectal tumors receiving 5-FU-based chemotherapy 2 tandem repeats was nearly twice as common in responders than in non-responders with longer median survival observed in patients with 2 tandem repeats compared to those with 3 tandem repeats (16 months vs. 12 months) (35).

Polymorphisms of TS that result in 5-FU variability of its pharmacodynamic and pharmacokinetic properties illustrate the important fact that cancer being a polygenic disease will need a polygenic solution to guide therapy. To accurately select patients that are most likely to tolerate and respond to 5-FU therapy, one needs to use combined genotyping of DPYD and TSER functional variants as well as taking into account nongenetic factors.

**TAXANES**

Paclitaxel and Docetaxel are the two most commonly used taxanes in the treatment of both early and advanced stage breast cancer (36). They exert their cytotoxic effects by binding to B-tubulin that subsequently results in the stabilization of microtubules and disruption of mitotic spindle formation during cell division ultimately resulting in cell death. Genes involved in drug transport, metabolism, and drug target all play an important role in overall taxane efficacy, with their polymorphisms partly explaining the variability observed in toxicity and response to these drugs.

The multidrug-resistant transporter, P-glycoprotein, encoded by ABCB1 (MDR1) and expressed in multiple cell types, has been shown to mediate both paclitaxel and docetaxel influx (37,38). Three polymorphisms in ABCB1 have been commonly studied (3435C>T, 1236C>Y, and 2677G>T/A) (39), with most studies showing inconsistent results when trying to establish an association between these polymorphisms and drug efficacy/toxicity. Studies focusing on looking at ABCB1 polymorphism in combination may yield more useful clinical information. Some of the recent studies have reported an association of the various ABCB1 polymorphisms with tumor response to paclitaxel (ABCB1
2677G>T/A) (40), risk of neutropenia with docetaxel (ABCB1 3435C>T) (41), and risk of neuropathy with paclitaxel (ABCB1 3435C>T) (42). Oxidative taxane metabolism occurs via the CYP450 pathway, with CYP2C8 and CYP3A4 primarily responsible for paclitaxel metabolism while docetaxel is mainly metabolized by CYP3A4 (43,44). The functional significance of the polymorphisms of these enzymes has not been established with associations seen in vivo studies (45) not observed in studies of cancer patients (46).

TAMOXIFEN

Tamoxifen is a selective estrogen receptor modulator (SERM) that is widely used in the treatment and prevention of breast cancer. When given in the adjuvant setting, it reduces the annual risk of recurrence of breast cancer by half and the breast cancer mortality rate by one-third (47). The pharmacology and metabolism of tamoxifen is complex and undergoes extensive phase I and II metabolism with the production of endoxifen and to a lesser extent 4-hydroxytamoxifen (48). These metabolites play an important role in the anticancer effect of tamoxifen through their high-affinity binding to estrogen receptors and suppression of estradiol-stimulated cell proliferation (49).

The clinical effects of tamoxifen are highly variable (50) due to the variability observed in the plasma levels of the more potent metabolites of tamoxifen. Tamoxifen is converted to endoxifen (4-hydroxy-N-desmethyltamoxifen) primarily by CYP2D6-mediated oxidation of N-desmethyl tamoxifen, while the conversion of tamoxifen to 4-hydroxy tamoxifen is catalyzed by multiple enzymes (51). Plasma concentrations of endoxifen is thus sensitive to variants of the CYP2D6 genotype with the most common variant making the enzyme inactive (common allele associated with CYP2D6 poor metabolizer phenotype is CYP2D6*4). One prospective trial revealed that women treated with tamoxifen who either carried the genetic variants associated with low or absent CYP2D6 activity or were treated concomitantly with drugs that inhibited the activity of CYP2D6 had lower levels of endoxifen (52). Subsequently, using tissue sample from a cohort of early stage breast cancer patients treated with tamoxifen, it was shown that women with a homozygous CYP2D6 phenotype tended to have a higher risk of disease relapse and a lower incidence of hot flashes (53).

Polymorphisms associated with the activity of the metabolites of tamoxifen teach us two important lessons. First, the interindividual variability in the response to tamoxifen is explained both by genetic variation in CYP2D6 and the coadministration of drugs that inhibit the activity of CYP2D6 had lower levels of endoxifen (52). Subsequently, using tissue sample from a cohort of early stage breast cancer patients treated with tamoxifen, it was shown that women with a homozygous CYP2D6 phenotype tended to have a higher risk of disease relapse and a lower incidence of hot flashes (53).

Second, the toxicity associated with drug can help identify responders.

CONCLUSION

At present time, dosages of anticancer drugs are administered based on the assumption of homogeneity among individuals and are rarely based on individual characteristics. As a result we encounter a situation where by a large
In this chapter, we will explore the interplay between pharmacology, pharmacogenetics, and pharmacoepidemiology. A significant number of individuals experience unwanted adverse side effects after treatment without attaining maximum drug efficacy. Such adverse drug reactions impose a huge burden not only on the patient but also on the health care system. A better understanding of the pharmacoepidemiological, pharmacological, and pharmacogenetic determinants of a chemotherapeutic agent is needed to optimize drug dosing regimens to obtain maximal efficacy with minimal toxicity. Advancements in the field of pharmacogenetics hold promise in helping to identify which polymorphisms are likely to be important in clinical oncology thereby enabling the development of effective agents that will make their way to the clinic in a timely manner. Realizing the importance of each of the components of individualized therapy and incorporating each one into the infrastructure of clinical trials will allow us to further explore and define the relationships between different polymorphisms and observed phenotypes. Considering the fact that anticancer drugs have narrow a therapeutic range oncology is a field that will certainly benefit from the application of pharmacogenetic and pharmacological principles to dosage individualization. The idea of the “one model fits all” should be accepted as grossly flawed and, as we progress in the field of cancer therapeutics, will be replaced in the future with the application of “personalized cancer chemotherapy” into routine clinical practice.

REFERENCES


Role of Genetic Variability in Breast Cancer Treatment Outcomes

Kandace L. Amend
Department of Oncological Sciences, Mount Sinai School of Medicine,
New York, New York, U.S.A.

Ji-Yeob Choi and Christine B. Ambrosone
Department of Cancer Prevention and Control, Roswell Park
Cancer Institute, Buffalo, New York, U.S.A.

INTRODUCTION

Variability in toxicity and response to the same dose of medication may occur among cancer patients and contribute to the unpredictability of clinical outcomes. In addition to the potential importance of clinical factors in determining drug effects (1), inherited differences in drug metabolism are likely to have a profound effect on treatment outcomes (2). Pharmacogenetics, the study of the contribution of genetic variation to these interindividual differences, may allow for the individualization of therapy and improved prediction of pharmacodynamic events, including host toxicity and treatment efficacy. To date, the role of pharmacogenetics in breast cancer therapy has not been clearly defined. Herein we discuss the drugs most commonly used to treat breast cancer and describe their metabolic pathways, as well as provide epidemiologic support for the role of pharmacogenetics in breast cancer outcomes.
BREAST CANCER THERAPEUTICS

The most frequently used agents to treat breast cancer include cyclophosphamide (CP) in combination with doxorubicin [Adriamycin\textsuperscript{H} (A)] or with methotrexate (MTX) and 5-fluourouracil (5FU) (3). The taxanes paclitaxel (Taxol\textsuperscript{H}) and docetaxel (Taxotere\textsuperscript{H}) have an important role in the treatment of early-stage breast cancer as well as metastatic disease. In patients with estrogen receptor (ER)-positive tumors, tamoxifen (TAM) reduces the risk of recurrence by 31% and breast cancer mortality by one-third (4), and randomized trials have indicated that treatment with third-generation aromatase inhibitors (AIs) is now an important option for the adjuvant treatment of hormone receptor–positive breast cancer (5,6).

Multiple enzymatic pathways are involved in the metabolism and detoxification of chemotherapeutic agents, and common genetic polymorphisms in drug metabolizing enzymes are likely to contribute to interindividual variability in toxicity and cancer recurrence (7).

Cyclophosphamide

Cyclophosphamide requires multistep activation before it can function as an antitumor alkylating agent. The drug is first oxidized in the liver to the active 4-hydroxy metabolite mainly by CYP2B6, although other enzymes, including CYP2A6, CYP3A4, CYP2C8, CYPC29, and CYP2C19, are involved. Drug intermediates are transported in blood and decompose spontaneously to phosphoramide mustard and acrolein, both of which can react with nucleophilic groups, forming DNA cross-links and inhibition of protein synthesis (8). Inactivating glutathionyl conjugation reactions of C intermediates are catalyzed by glutathione S-transferases (GSTs) GSTA1 and GSTP1 (9). Variability in the activity of these activating and detoxifying enzymes is likely to affect ultimate tumor tissue dose and efficacy of treatment.

The human cytochrome P450 CYP2B6 gene is involved in the metabolic activation of a number of clinically important chemotherapeutic drugs for breast cancer, including CP (10,11) as well as the antioestrogen TAM (12). Several single nucleotide polymorphisms (SNPs) in CYP2B6 have been described, some having functional significance (13) and affecting the pharmacokinetic parameters of CP. In cell culture, enzymatic activities in microsomes from COS-1 cells expressing a K262R amino acid substitution (CYP2B6*4, CYP2B6*6, and CYP2B6*7) showed increased values for $V_{\text{max}}$ and $V_{\text{max}}/K_{\text{m}}$ compared with that of the wild type (CYP2B6*1) (14), while a second study reported that the amino acid substitution Q172H resulting from the CYP2B6*6 allele enhanced 7-ethoxycoumarin O-deethylase activity (15). A study by Lang et al. (2001) identified a total of nine SNPs by sequencing DNA from 35 subjects and reported extensive variability in the expression and activity of CYP2B6, where CYP2B6 expression was significantly reduced in carriers of the R487T, Q172H,
Role of Genetic Variability in Breast Cancer Treatment Outcomes

and K262R polymorphisms (16). Ethnic variation in allelic variants and CYP2B6 expression has been also been observed (17,18).

CYP3A4 also plays a large role in the metabolism of CP, and the CYP3A4 gene has a number of known variant alleles, including some having functional significance (19). A single base substitution in the 5′ promoter region of the gene (20), present in 9% of European-Americans and 53% of African-Americans (21), with the CYP3A4*1B allele associated with expression levels is 1.4-fold higher than that in common alleles (22). Using enzyme kinetic analyses to evaluate intersample variation in activation of CP by CYP3A and 2B in human liver, both Chang et al. (1993) and Roy et al. (1999) found considerable interindividual differences in CYP3A and CYP2B levels and variation over a wide range in CP hydroxylation (10,23). Variation in CYP3A enzyme levels and expression may be related, in part, to genetic differences or to other chemotherapeutic and natural agents that have been shown to increase CYP3A4 expression (24).

CYP3A5 accounts for at least 50% of the total CYP3A content in most Caucasians and African-Americans and is thus considered an important contributor to interindividual differences in CYP3A-dependent drug clearance (25). Although CYP3A5 is polymorphically expressed at high levels in a minority of Caucasians, the expression in human liver is more frequent in African-Americans (60%) than in Caucasians (33%) (25). Reduced expression in some individuals may be due to a common polymorphism in intron 3 of CYP3A5 (A6986G), which generates a cryptic splice site (CYP3A*6) resulting in a truncated amino acid (25,26).

Genetic polymorphisms have been described in several of the CYP2C family members, including CYP2C8, CYP2C29, CYP2C18, and CYP2C19, although the catalytic contribution of CYP2C8 and CYP2C18 to activation of CP is likely minor compared with the contributions of CYP2C9 and CYP2C19. In vitro, CYP2C19 was found to bioactivate CP in human liver microsomes (27), and while both CYP2C9 and CYP2C19 were found to possess CP 4-hydroxylase activity in vitro, CYP2C19 exhibited higher CP 4-hydroxylase activity following only CYP2B6 (23). Although known variants in CYP2C9 and CYP2C19 exist, little is known regarding their contribution to pharmacokinetic variability in relation to breast cancer prognosis. In a study by Xie et al., CYP2C9*2, CYP2C9*3, CYP2C19*2, and CYP2C19*3 were not significantly correlated with the metabolism of CP in 19 human liver specimens (11).

Phase II drug metabolism is catalyzed by GSTs, which consist of several isoenzymes, including mu (GSTM1), pi (GSTP1), alpha (GSTA1), theta (GSTT1), and zeta. The GSTs detoxify many chemotherapeutic drugs, including CP, are induced by oxidative stress, and detoxify the reactive oxygen species (ROS) that are produced by chemotherapeutic drugs and radiation therapy. GSTA1, the enzyme most active in glutathione conjugation reactions with CP intermediates, has a polymorphism in the 5′ promoter region of the gene (28). Coles et al. (2001) identified substitutions designated as GSTA1*A (−567, −69, and −52) and GSTA1*B (−631, −567, and −69) (29), with the GSTA1*B variant
showing decreased hepatic expression and increased GSTA2 expression compared with GSTA1*A homozygotes. Given that GSTA1 increases the conjugation of CP intermediates (9), these differences in hepatic expression may affect the detoxification of CP metabolites and, ultimately, the outcomes of treatment and survival.

For the GSTP1 gene, two variant alleles, GSTP1*B and GSTP1*C, have been detected (30). A single amino acid change from isoleucine to valine occurs at codon 105 (Ile105Val) of the GSTP1 gene, with genotype frequencies in Caucasians 51% homozygous for Ile/Ile, 43% heterozygous for Ile/Val, and 6% homozygous for the variant allele Val/Val (31). The allelic variant 105Val exhibited lower glutathione conjugation of the alkylating agent thiopeta compared with the isoleucine allele (32). GSTP1 variants may also differ in activity toward CP, structurally similar to thiotepa, and patients with GSTP1 allelic variants may have reduced activity in removal of active CP.

**Anthracyclines**

The commonly used anthracyclines include doxorubicin, daunorubicin, and epirubicin, although doxorubicin and epirubicin are used more frequently for the treatment of breast cancer. Several different mechanisms have been proposed for anthracycline tumor activity, including free radical generation, DNA binding and alkylation, DNA cross-linking, interference with DNA unwinding or strand separation, direct membrane effects, inhibition of topoisomerase II, and induction of apoptosis (33). Doxorubicin, like other anthracyclines, results in the formation of quinone-mediated free radicals, which have the capacity to cause oxidative damage and cytotoxicity, resulting in significant cardiotoxicity as a result of the production of ROS (34). Cellular oxidoreductases reduce Adriamycin to a semiquinone radical that is subsequently reoxidized by oxygen to a superoxide anion and the parent quinine (35–37). Superoxide anions can dismutate to form hydrogen peroxide and/or react with nitric oxide to form peroxynitrite. Lipid peroxides resulting from doxorubicin can further break down to yield hydroxyalkenals, which are substrates for glutathione-conjugating isozymes (38), indicating the potential importance of GSTs in mediating therapeutic outcomes.

Carbonyl reductase (CBR) is a monomeric, NADPH-dependent oxidoreductase that catalyzes the reduction of a variety of carbonyl compounds, including doxorubicin and daunorubicin. In humans, two CBRs have been identified, CBR1 and CBR3. An SNP has been identified in CBR3 (CBR3 V244M), where the M244 isoform results in higher $V_{\text{max}}$ and is 100% more efficient in catalyzing the reduction of substrate (39). To our knowledge, CBRs have not been investigated in relation to breast cancer outcomes following treatment with anthracyclines.

In addition to its mechanistic role in cytotoxicity of anthracyclines, it is well documented that oxidative stress plays a role in the cell-killing abilities of many
other chemotherapy agents, as well as in radiation therapy. Specifically for breast cancer, CP, doxorubicin (Adriamycin), and paclitaxel and radiation therapy generate ROS and induce apoptosis, which may contribute to treatment efficacy and increased survival. ROS can cause mitochondrial permeability, and once the mitochondrial membrane barrier function is lost, a number of other factors contribute to cell death. While ROS, among other factors, induce or facilitate mitochondrial permeability, glutathione and antioxidant enzymes such as manganese superoxide dismutase (MnSOD), catalase (CAT), and glutathione peroxidase (GPX1) inhibit it (40). In the mitochondrion, MnSOD catalyzes the dismutation of two superoxide radicals, producing H$_2$O$_2$ and oxygen. If H$_2$O$_2$ is not neutralized, it may contribute to further generation of ROS by a reaction catalyzed by myeloperoxidase (MPO). Thus, levels of MnSOD, CAT, GPX1, and MPO are important in determining the amount of ROS or detoxification of H$_2$O$_2$. It is plausible that interindividual variability resulting from polymorphisms in these genes may influence the efficacy of cancer treatment and survival.

Two genetic polymorphisms appear to affect mitochondrial levels of MnSOD. A single-base substitution variant in the *MnSOD* gene results in an amino acid substitution from alanine to valine at position 9 (ala-9) in the mitochondrial targeting sequence (41). The substitution alters the secondary structure of the protein, affecting the cellular distribution of the enzyme and transport of MnSOD in the mitochondrion, where it would be biologically available (41). In the mitochondria, the *MnSOD* Ala variant has been shown to generate more active *MnSOD* protein compared with the Val allele (42). A second polymorphic variant has been identified in *MnSOD*, and Ile or Thr at amino acid position 58 (43) and the Thr$^{58}$MnSOD variant was shown to exhibit only half the enzymatic activity of Ile$^{58}$MnSOD (43,44). It can be hypothesized that the variant (ala-9 and Thr$^{58}$MnSOD) would confer reduced protection of breast tumor cells from ROS produced during therapy. Thus, lesser MnSOD activity due to polymorphisms could increase treatment efficacy but also result in greater toxicity due to ROS.

CAT is most abundant in the liver, kidneys, and erythrocytes, where it decomposes H$_2$O$_2$ to H$_2$O and O$_2$. A common functional polymorphism (C-262T) alters transcription factor binding and the expression of CAT in erythrocytes. We previously found that individuals with C alleles had significantly higher CAT levels compared with those with T/T genotypes (45).

H$_2$O$_2$, if not neutralized, may contribute to further generation of ROS by a reaction catalyzed by MPO. MPO generates ROS endogenously by functioning as an antimicrobial enzyme, catalyzing a reaction between H$_2$O$_2$ and chloride to generate hypochlorous acid (HOCI), a potent oxidizing agent. HOCI further reacts with other biological molecules to generate secondary radicals (46). The promoter region of the *MPO* gene has a G-to-A base substitution at position 463, which creates high- (G) and low-expression (A) alleles (47). Variability in genes in this oxidative stress-related pathway may influence ultimate levels of ROS and, thereby, tumor cell kill.
5-Fluorouracil and Methotrexate

The agents 5FU and MTX are involved in folate metabolism, in the conversion of homocysteine to methionine, and in purine and pyrimidine synthesis. The activated metabolites of 5FU, similar to pyrimidine nucleotides, competitively and reversibly bind fluoropyrimidine thymidylase synthase (TS) inhibitors, forming covalently bound complexes with the TS enzyme, thus inhibiting the enzyme’s activity (48). Folate antagonists, such as MTX, are structurally similar to folate and inhibit the action of various enzymes such as dihydrofolate reductase (DHFR) (49). Their action is accomplished through incorporation as false precursors in DNA or RNA or through inhibition of proteins involved in nucleotide metabolism.

Common polymorphisms have been reported in many enzymes and transport proteins involved in intracellular folate metabolism. Using a candidate gene approach, pharmacogenetic research, to date, has focused on TS because of its role as a drug target, dihydropyrimidine dehydrogenase (DPD) because of its involvement in pyrimidine degradation, and methylenetetrahydrofolate reductase (MTHFR) because of its key involvement in the control of the flux of intracellular folate metabolites (50). A genetic polymorphism in the 5' regulatory region of the TS gene promoter, consisting of either double (2R) or triple (3R) repeats of a 28-bp sequence (51), has been shown to influence TS expression, with higher expression in 3R/3R tumors (52). A G>C polymorphism has been identified within the second repeat of 3R alleles that alters the transcriptional activation of the TS gene constructs bearing this genotype (53). A 6-bp deletion (1494del6) within the 3'UTR (54) has been also associated with decreased mRNA stability (55). Several clinical studies have shown the potential contribution of the above TS germinal polymorphism in the prediction of tumor responsiveness and/or toxicity following treatment by fluoropyrimidines (56). However, the role of TS polymorphism as a predictor of treatment outcome is still unclear.

MTHFR catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which is an essential cofactor in the biosynthesis of deoxythymidine monophosphate. A common C677T transition in exon 4 and an A1298C transition in exon 7 of the MTHFR gene result in lower specific activity (57,58). Most studies, to date, have examined the association between 5FU and MTHFR polymorphisms in the context of colorectal cancer, but Toffoli et al. (59) reported that of six patients who developed severe acute toxicity in the first cycle of adjuvant CMF, five had the MTHFR 677 T/T genotype and one had the C/C genotype.

Because 5FU has a relatively narrow therapeutic index, toxicity increases as the dose is increased, and more than 80% of administered 5FU is catabolized by DPD (1). Patients with complete or partial deficiency of DPD often experience severe or even life-threatening toxicity after the administration of 5FU (60). At least 11 of 33 mutations identified in the dihydropyrimidine dehydrogenase (DPYD) gene have been detected in patients suffering from severe 5FU-associated toxicity (61–63).
There have been few studies of \textit{MTHFR} and \textit{TS} polymorphisms in relation to breast cancer treatment outcomes. In a study with 19 human cancer cell lines, including breast cancer, Etienne et al. (64) found that there was a marked trend for greater 5FU efficacy in mutated \textit{MTHFR} 1298C cell lines compared with those with A/A genotypes ($P = 0.055$), but \textit{MTHFR} C677T and TS 2R/3R polymorphisms were not associated with 5FU sensitivity. Largillier et al. (65) noted that genetic polymorphisms of \textit{MTHFR} (677C>T and 1298A>C), \textit{DPD} (IVS14+1G>A), and \textit{TS} were associated with efficacy and toxicity associated with capecitabine, an oral 5FU prodrug, in 105 breast cancer patients receiving capecitabine as monotherapy.

Although MTX has been widely used to treat many cancers, including leukemia, lymphomas, and breast cancer, as well as some autoimmune diseases, there has been limited research to evaluate the effects of genetic variability on treatment outcomes. Associations have been noted for polymorphisms in folate-related genes in leukemia (66–69), and Sohn et al. (70) found that in vitro transfection of mutant 677T MTHFR cDNA in breast and colon cancer cells increased the chemosensitivity of these cells to 5FU but decreased the chemosensitivity of breast cancer cells to MTX, suggesting at least a modulating role of this genotype on cellular fluoropyrimidine sensitivity.

\textbf{Taxanes}

The taxanes paclitaxel and docetaxel interfere with microtubular disassembly, ultimately resulting in DNA fragmentation and features of apoptosis (71). The contributions of genetic variability to paclitaxel metabolism and breast cancer outcomes is still unclear. CYP2C8 is thought to be the main CYP-metabolizing enzyme of paclitaxel, while CYP3A4 and CYP3A5 play more minor roles in drug metabolism (19). At least three known variants in \textit{CYP2C8} exist, and the \textit{CYP2C8*3} variant allele has been associated with decreased paclitaxel 6-\textit{a}hydroxylase activity in human cell lines and human liver microsomes (72). In a recent study of paclitaxel metabolism in breast cancer, however, \textit{CYP2C8} genotypes (\textit{CYP2C8*3} and \textit{CYP2C8*4}) were not significantly associated with paclitaxel clearance (73). Ethnic variation exists in the distribution of \textit{CYP2C8} alleles, with the \textit{CYP2C8*2} allele more common in African-Americans than Caucasians (18% vs. 0%, respectively) and the \textit{CYP2C8*3} allele more common in whites (13%) than African-Americans (2%). The \textit{CYP2C8*5} allele is found in 9% of Asians (74).

\textbf{Antiestrogens}

The pharmacology and metabolism of TAM is complex, and differences in patient outcomes could be due to individual variation in the metabolism of the drug. TAM is metabolized primarily by CYP3A4, CYP2D6, CYP2C9, SULT1A1, and UGT2B15 (75–78). The main metabolites of TAM are \textit{N}-desmethytamoxifen (formed by CYP3A4) and endoxifen and 4-OH-TAM.
CYP2D6 is polymorphic with 46 known allelic variants, which are categorized as resulting in abolished, decreased, normal, increased, or qualitatively altered catalytic activity (79). The major polymorphic variant CYP2D6 alleles include CYP2D6*2, CYP2D6*4, CYP2D6*5, CYP2D6*10, CYP2D6*17, and CYP2D6*41 (79).

Human sulfotransferase 1A1 (SULT1A1) catalyzes the sulfation of a variety of phenolic and estrogenic compounds, including endogenous and environmental estrogens and the 4-hydroxy metabolite of TAM, 4-OH TAM (80). An exon 7 polymorphism in SULT1A1 (SULT1A1*2) results in a translated protein with approximately twofold lower catalytic activity and decreased thermostability than the common allele (SULT1A1*1) (81). We found, as hypothesized, that women treated with TAM who were homozygous for the variant SULT1A1*2/*2 allele had approximately three times the risk of death compared with those homozygous for the common allele or heterozygotes (82).

A wide range of drugs are substrates for glucuronosyltransferases (UGTs) and this pathway has been estimated to account for approximately one-third of all drugs metabolized by phase II drug metabolizing enzymes (2). SULT1A1 and UGT2B15 have overlapping substrate specificity for 4-OH TAM. A functional polymorphism in UGT2B15, a G→T substitution at codon 85 within the putative substrate recognition site of the gene, results in higher V_max compared with the common allele (83). In breast cancer patients treated with TAM, the combined effects of SULT1A1 and UGT2B15 were evaluated, and an increasing number of variant SULT1A1 and UGT2B15 alleles was associated with poorer survival (84).

Aromatase Inhibitors

AIs, including exemestane, anastrozole, and letrozole, have come into clinical use more recently, developed primarily for hormone-dependent breast cancer in postmenopausal women. AIs act by inhibiting the cytochrome P450 enzyme, aromatase (CYP19), which catalyzes the conversion of androgens to estrogens. AIs block this enzyme, decreasing circulating levels of estrogens. It has been demonstrated that exemestane is metabolized by CYP3A4 but is also metabolized by aldoketoreductases (85). Letrozole is metabolized by CYP2A6 and CYP3A4 to a pharmacologically inactive carbinol metabolite (86), and anastrozole is metabolized by N-dealkylation and hydroxylation reactions followed by glucuronidation. However, specific isoforms involved and the effects on drug levels have not yet been identified.

MOLECULAR EPIDEMIOLOGIC STUDIES OF PHARMACOGENETICS AND BREAST CANCER OUTCOMES

To date, most analytical studies of pharmacogenetics and breast cancer outcomes have been conducted among patients receiving heterogeneous chemotherapy treatments, although most chemotherapy regimens included CP. Several studies
Role of Genetic Variability in Breast Cancer Treatment Outcomes

have evaluated polymorphisms in phase II or GST enzymes in relation to prognosis after treatment for breast cancer. In a study of 92 women with advanced breast cancer, the GSTM1-null genotype was not significantly associated with tumor characteristics, disease-free survival (DFS), or overall survival (OS) (87), but in a retrospective study with a larger sample size, we found that women with low-activity homozygous genotypes for the GSTA1*B allele had better survival, with the hazard of death reduced by 30% (88,89), and women with low-activity GSTP1 Val/Val genotype had a similar risk reduction (89). Deletion polymorphisms in GSTM1 and GSTT1, genes responsive to ROS and resulting from radiation or chemotherapy and subsequent lipid peroxidation, were associated with a reduced hazard of death compared with women with both alleles present [HR = 0.59; 95% confidence interval (CI), 0.36–0.97; and HR = 0.51; 95% confidence interval (CI), 0.29–0.90] (90). In a study in China (91), women with GSTP1 Val/Val genotypes experienced a 60% reduction in overall mortality (HR = 0.4; 95% confidence interval (CI), 0.2–0.8), although no associations were found with GSTM1 or GSTT1 genotypes. A smaller study (n = 85) evaluated polymorphisms in both phase I and phase II metabolism of CP among women with metastatic or inflammatory breast cancer treated with high-dose CP, cisplatin, and carbustine. Patients with the GSTM1-null genotype had prolonged OS, while those with variants in CYP3A4*1B and CYP3A5*1 had shortened survival. To our knowledge, only one genotype association study has focused on paclitaxel metabolism. In that study of 93 patients (73), there were no significant associations between ABCB1, ABCG2, CYP1B1, CYP3A4, CYP3A5, and CYP2C8 and paclitaxel clearance. However, patients homozygous for CYP1B1*3 Leu/Leu genotype had significantly longer progression-free survival (PFS) compared with patients with at least one Val allele (Table 1).

Few studies have examined the effects of MTHFR genotypes on breast cancer survival. In the Shanghai Breast Cancer Study (92), no significant associations were found between MTHFR genotypes and the risk of death. A second study evaluated these same genotypes in African-American and Caucasian women from the Baltimore area (93) and found that carriers of the low-activity allele at codon 1298 (A/C or C/C) had significantly poorer survival, while those with the variant allele at codon 677 (C/T or T/T) had improved survival. Associations were stronger in patients with ER-negative tumors, and African-American women had improved survival if they were carriers of either variant allele (A1298C, P_interaction = 0.088, C677T, P_interaction = 0.026).

Molecular epidemiologic studies have indicated that genetic variation in phases I and II conjugating enzymes can influence the efficacy of TAM therapy for breast cancer. In a recent study in Sweden (94), patients homozygous for CYP2D6*4 alleles experienced significantly better DFS compared with patients with at least *1 allele (P = 0.05). In a group of women randomized to 5, rather than two years of TAM, patients homozygous for CYP3A5*3 experienced improved relapse-free survival (RFS) (HR = 0.02; CI, 0.07–0.55). However, this finding should be interpreted with caution since the number of patients within
### Table 1 Epidemiologic Studies of Genetic Variation in Chemotherapeutic Metabolizing Genes and Breast Cancer Outcomes

<table>
<thead>
<tr>
<th>Genes</th>
<th>SNP</th>
<th>Participants</th>
<th>Treatment</th>
<th>Outcome</th>
<th>Risk (95% CI)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclophosphamide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTM1</td>
<td>GSTM1-null genotype</td>
<td>92 patients</td>
<td>CAF</td>
<td>OS or DFS</td>
<td>GSTM1 not significant</td>
<td>87</td>
</tr>
<tr>
<td>GSTA1</td>
<td>GSTA1<em>B/B</em></td>
<td>245 patients</td>
<td>CP based</td>
<td>5-yr survival</td>
<td>GSTA1<em>B/B</em> HR = 0.3 (0.1–0.8)</td>
<td>88</td>
</tr>
<tr>
<td>GSTP1</td>
<td>GSTP Val/Val</td>
<td>240 patients</td>
<td>CP based</td>
<td>OS</td>
<td>GSTP Val/Val HR = 0.3 (0.1–1.0)</td>
<td>89</td>
</tr>
<tr>
<td>GSTP1</td>
<td>GSTP Val/Val</td>
<td>1034 patients</td>
<td>CP based</td>
<td>OS</td>
<td>GSTP Val/Val HR = 0.4 (0.2–0.8)</td>
<td>91</td>
</tr>
<tr>
<td>6 genes, CYPs,</td>
<td>8 SNPs</td>
<td>93 patients</td>
<td>High-dose paclitaxel and CP and A</td>
<td>PFS</td>
<td>CYP1B1*3 Leu/Leu Longer PFS (P = 0.037)</td>
<td>73</td>
</tr>
<tr>
<td>ABCB1, ABCG2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 genes, CYPs,</td>
<td>29 SNPs</td>
<td>85 patients</td>
<td>CP and cisplatin and carmustine</td>
<td>PK of CP and OS</td>
<td>GSTM1-null genotype; better OS; at least one variant allele in CYP3A4<em>1B or CYP3A5</em>1</td>
<td>(98)</td>
</tr>
<tr>
<td>METs, GSTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methotrexate and 5-fluorouracil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTHFR</td>
<td>MTHFR C677T; MTHFR A1298C</td>
<td>1067 patients</td>
<td>Combination (most received 5FU)</td>
<td>OS</td>
<td>MTHFR genotypes not significant with OS</td>
<td>92</td>
</tr>
<tr>
<td>MTHFR</td>
<td>MTHFR C677T; MTHFR A1298C</td>
<td>248 patients</td>
<td>Combination (most received 5FU)</td>
<td>OS</td>
<td>MTHFR 1298 A/C or C/C HR = 2.05 (1.05–4.0); MTHFR 677 CT or T/T HR = 0.65 (0.31–1.35)</td>
<td>(93)</td>
</tr>
<tr>
<td>Tamoxifen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP3A5; CYP2D6; SULT1A1;</td>
<td>CYP3A5<em>3; CYP2D6</em>4; SULT1A1*2; UGT2B15</td>
<td>677 patients</td>
<td>TAM</td>
<td>RFS</td>
<td>CYP3A5 *3/*3 in 5-year TAM group HR = 0.20 (0.07–0.55); CYP2D6 *4/*4 better RFS (n = 677) (P = 0.05)</td>
<td>(94)</td>
</tr>
<tr>
<td>UGT2B15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP3A5; CYP2D6</td>
<td>CYP3A5<em>3; CYP2D6</em>4;</td>
<td>223 patients</td>
<td>TAM</td>
<td>RFS or DFS or OS</td>
<td></td>
<td>(95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 1: Epidemiologic Studies of Genetic Variation in Chemotherapeutic Metabolizing Genes and Breast Cancer Outcomes (Continued)

<table>
<thead>
<tr>
<th>Genes SNP Participants Treatment Outcome</th>
<th>Risk (95% CI) Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2D6<em>4; UGT2B15</em>2; SULT1A1*2</td>
<td>84</td>
</tr>
<tr>
<td>337 patients (162 patients treated with TAM) TAM OS or PFS</td>
<td>CYP2D6 NS; UGT2B15<em>2 and SULT1A1</em>2 HR = 4.4 (1.17–16.55) in TAM group</td>
</tr>
<tr>
<td>SULT1A1*2</td>
<td>(82)</td>
</tr>
<tr>
<td>377 patients (160 treated with TAM) TAM OS SULT1A1*2/*2</td>
<td>HR = 2.9 (1.1–7.6) in TAM group</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oxidative stress</th>
<th>GSTM1/GSTT1</th>
<th>MPO, MnSOD, CAT</th>
<th>MnSOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTM1/GSTT1</td>
<td>GSTM1 and GSTT1-null genotypes</td>
<td>Combination OS and DFS GSTM1-null genotype HR = 0.59 (0.36–0.97); GSTT1-null genotype HR = 0.51 (0.29–0.90)</td>
<td>90</td>
</tr>
<tr>
<td>MPO, MnSOD, CAT</td>
<td>MnSOD Ala-9Val, CAT-262C→T, MPO-463G→A</td>
<td>Combination OS MPO GG genotype HR = 0.60 (0.38–0.95), MnSOD CC genotype HR = 0.70 (0.40–1.24), genotypes combined HR = 0.33 (0.13–0.80)</td>
<td>96</td>
</tr>
<tr>
<td>MnSOD</td>
<td>MnSOD-102 C→T</td>
<td>Combination RFS MnSOD CT + TT HR = 0.65 (0.40–1.05) for chemotherapy group</td>
<td>97</td>
</tr>
</tbody>
</table>

Abbreviations: CAF, Cyclophosphamide/doxorubicin/fluorouracil; A, doxorubicin; CP, cyclophosphamide; T, DFS, disease-free survival; NS, not significant; OS, overall survival; PFS, progression-free survival; RFS, relapse-free survival; TAM, tamoxifen; SFU, 5-fluorouracil; GST, glutathione S-transferase; SNP, single nucleotide polymorphism; METs, Metallothioneins; MTHFR, methylenetetrahydrofolate reductase; MPO, myeloperoxidase; MnSOD, manganese superoxide dismutase; CAT, catalase; PT, Pharmacokinetics.
this group was small. No statistically significant differences were found for CYP2D6, SULT1A1, or UGT2B15 and the length of TAM treatment. In contrast, Goetz et al. (95) found no significant associations between CYP3A5*3 and breast cancer outcomes; however, women with the CYP2D6 *4/*4 genotype (n = 13) had statistically significant worse RFS (P = 0.023) and DFS (P = 0.12) but not OS (P = 0.169). However, after adjustment for nodal status and tumor size, associations were no longer statistically significant. These results are in agreement with our previous report, wherein we noted no significant associations between the CYP2D6*4 genotype and OS or PFS among breast cancer patients treated with TAM (84).

In addition to CYP isoforms, phase II enzymes, including sulfotransferase 1A1 (SULT1A1) and UDP-glucuronosyltransferase isoform 2B15 (UGT2B15), have been implicated in the metabolism of TAM, and genetic variability in these genes together may influence breast cancer survival. In a cohort of women with breast cancer who were treated with TAM, those who were homozygous for the low-activity SULT1A1*2 allele had a threefold higher risk of death compared with those with at least *1 common allele (HR = 2.9; 1.1–7.6), while no significant associations were found among women who did not receive TAM (82).

When considering the independent and combined effects of SULT1A1 and UGT2B15 genotypes on OS and recurrence among women treated with TAM, there was a nonsignificant trend toward increased risk of disease recurrence and poorer survival with increasing number of UGT2B15*2 alleles (P = 0.11) (84). However, there was a significant trend toward decreased OS among women treated with TAM, with increasing number of variant alleles when UGT2B15 and SULT1A1 genotypes were combined (P = 0.03), although these risk estimates may be unstable because only 16 cases had two variant alleles.

We previously found that genotypes for MnSOD and MPO resulting in higher levels of ROS were associated with better survival in a study of women from Arkansas, and the combined effects of MnSOD and MPO were greater than those for either polymorphism alone (HR = 0.33; CI, 0.13–0.80) (96). An additional MnSOD variant, located in the promoter region of the gene, was also associated with a significant decrease in RFS (CT+TT genotype HR = 0.65, CI, 0.40–1.05), although no significant associations were found with OS (97).

In the context of a clinical trial for breast cancer with CAF versus CMF (SWOG 8897) (unpublished data, presented at the San Antonio Breast Cancer Symposium, 2006), as in the Arkansas study, MPO genotypes encoding for higher activity were associated with better survival.

**FUTURE DIRECTIONS**

Studies, to date, indicate that interindividual variability in drug metabolism may affect treatment-related toxicities and DFS. However, much of the data are inconsistent, and it is likely that the evaluation of single genes in complex pathways presents only part of a complicated picture. Thus, future studies of
pharmacogenetics and treatment outcomes should focus on entire drug metabolism pathways in large sample sets receiving homogeneous treatment regimens. Furthermore, assessment of variability across an entire gene, through construction of haplotype blocks and tag SNPs, may give more complete information than a study of single SNPs. It is likely that, in the future, other, more systemic markers of inherited drug metabolism capabilities may better predict treatment outcomes, and scans of variability across the entire genome may reveal factors that are likely to determine treatment outcomes. The consideration of individual genotypes, as well as tumor characteristics, in therapeutic decision making holds promise for individualized treatment for breast cancer in the future.

REFERENCES


79. Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J 2005; 5(1):6–13.


The Role of Epigenetics in Breast Cancer: Implications for Diagnosis, Prognosis, and Treatment

Amy M. Dworkin, Tim H.-M. Huang, and Amanda E. Toland

Human Cancer Genetics Program, Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University, Columbus, Ohio, U.S.A.

INTRODUCTION

In the past 10 years, the field of epigenetics has expanded into an essential component of clinical cancer research and therapeutics. Epigenetic modifications in DNA or chromatin are stably transmitted over rounds of cell division but do not result in any permanent alterations to the underlying DNA sequence. These changes, such as histone modifications and DNA methylation of tumor suppressor genes, are considered a hallmark of certain cancers, including breast cancer. Unlike germline mutations, epigenetic modifications can potentially be reversed, making them very appealing for preventative care and therapeutics against cancers (1). Because of this, demethylating agents are currently being evaluated for efficacy in the treatment of breast cancer.

There are distinct mechanisms that initiate and sustain epigenetic modifications (2–8). Of these, DNA methylation and posttranslational modifications of histone proteins are the best understood. DNA methylation is a covalent addition of a methyl group to DNA, usually to a cytosine located 5' to guaninesine (CpG dinucleotide). CpG dinucleotides are underrepresented in the genome,
except for small clusters, referred to as CpG islands, located in or near the promoter of approximately half of all genes (1,9–11). In addition, other epigenetic modifications have recently been explored, including the Polycomb group (PcG) proteins, which repress gene function and can only be overcome by germline differentiation processes, and small noncoding RNA molecules, which regulate gene expression by targeting RNA degradation (3). Small noncoding RNA molecules have recently been found to also target gene promoters and result in transcriptional gene silencing (11,12).

DNA methylation plays a variety of roles in normal cells: silencing of transposable elements and viral sequences, maintenance of chromosomal integrity, X chromosome inactivation, and transcriptional regulation of genes (1). Methylation is established and maintained by at least four DNA methyltransferases (DNMTs), DNMT1, DNMT3A, DNMT3B, and DNMT3L, which catalyze the covalent addition of methyl groups to cytosine in the CpG dinucleotide. Methylation of CpG islands causes transcriptional inactivation (Fig. 1). These methyltransferases and their isoforms have been demonstrated to be essential for development; a knockout of any of these enzymes in mouse embryos results in embryonic or perinatal lethality (1). Functionally, there are two types of methyltransferases, maintenance and de novo. Maintenance methyltransferases are essential for copying DNA methylation patterns during cell division, and de novo methyltransferases introduce new patterns of methylation during early development. There are duplications of functions between these methylases so that perturbation of both types may be necessary to change overall methylation patterns (6). DNMT1 functions as a maintenance methyltransferase by copying DNA methylation patterns to daughter strands, following replication. It is the most abundant and catalytically active form. DNMT3A and DNMT3B act on both unmethylated and hemimethylated DNA. However, they are thought to have different methylation targets depending on the cell type and stage of development (13). In vitro knockouts of DNMT3a or DNMT3b each exhibit distinct phenotypes. DNMT3a-null cells lack methylation at imprinting loci, while DNMT3b-null cells result in loss of DNA methylation on endogenous
C-type retroviral DNA. Deletion of both DNMT3a and DNMT3b abolishes de novo methylation, while DNMT1 deletion produces bulk DNA demethylation (7). Of the methyltransferases, DNMT3L is unique. It stimulates de novo methylation by DNMT3a, does not enhance its binding to DNA, and alone does not bind to DNA (14).

Along with methylation, posttranslational modification of histones plays an important role in epigenetics. Histone proteins associate with DNA to form nucleosomes, permitting large amounts of DNA to be neatly packed into the nucleus. There are two configurations of chromatin, heterochromatin and euchromatin. Heterochromatin is condensed and transcriptionally inactive, while euchromatin has an “open” configuration and is favorable for gene transcription (Fig. 2). The N-terminal tails of histones “stick out” of the nucleosome and are subject to posttranslational modification such as acetylation, phosphorylation, methylation, ubiquitination, and sumoylation (2,4,7,15). The pattern of histone modification creates a “code,” which is read by proteins involved in chromatin remodeling and the dynamics of gene transcription (6,11). Acetylation and deacetylation are the most common types of histone modification. Histone deacetylases (HDACs) are a class of enzymes that remove acetyl groups from an ε-N-acetyl lysine amino acid on a histone. Its action is opposite to that of histone acetyltransferases (HATs). Deacetylation removes acetyl groups from histone tails, causing the histones to wrap more tightly around the DNA and interfering with transcription by blocking access to transcription factors. The overall result of histone deacetylation is a global (nonspecific) reduction in gene expression (Table 1).

**METHYLATION IN CANCER**

Cancer development is complex due to progressive changes of the genome that result in altering normal cellular processes such as cell growth, apoptosis, and angiogenesis (12). Both genetic and epigenetic modifications play a strong role in tumorigenesis (7). In classic tumor biology, cells lose both copies of tumor
Table 1  Histone Modifiers

<table>
<thead>
<tr>
<th>Histone modifier</th>
<th>Mechanism</th>
<th>Catalysis</th>
<th>Role in transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>Introduction of the acetyl group to the lysine residue at the N terminus of the histone protein</td>
<td>HATs</td>
<td>Enhances transcription (removes positive charge and reduces affinity between histones and DNA)</td>
</tr>
<tr>
<td>Deacetylation</td>
<td>Removal of the acetyl group from the lysine residue at the N terminus of the histone protein</td>
<td>HDACs</td>
<td>Represses transcription</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>Addition of phosphate group</td>
<td>Histone kinases</td>
<td>Enhances transcription (neutralizes histones’ basic charge and reduces affinity to DNA)</td>
</tr>
<tr>
<td>Dephosphorylation</td>
<td>Removal of phosphate group</td>
<td>Histone phosphatases</td>
<td>Represses transcription</td>
</tr>
<tr>
<td>Methylation</td>
<td>Addition of 1, 2, or 3 methyl groups on either lysine or arginine residues</td>
<td>HMTs</td>
<td>Silences genes</td>
</tr>
<tr>
<td>Demethylation</td>
<td>Removal of methyl groups</td>
<td>Cytochrome P450 family of liver enzymes</td>
<td>Activates genes</td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>Covalent attachment of ubiquitin</td>
<td>Enzymatic cascade, including ubiquitin ligase, E1 ubiquitin activating enzyme, and ATP</td>
<td>Labels protein for degradation</td>
</tr>
<tr>
<td>Deubiquitination</td>
<td>Removal of ubiquitin</td>
<td>Isopeptidases</td>
<td>Reverses ubiquitination</td>
</tr>
<tr>
<td>Sumoylation</td>
<td>Covalent attachment of SUMO to lysine residues</td>
<td>Enzymatic cascade, including SUMO-activating enzyme and ATP</td>
<td>Can mediate gene silencing through recruitment of HDAC and heterochromatin protein 1</td>
</tr>
</tbody>
</table>

Abbreviations: HATs, histone acetyltransferases; HDACs, histone deacetylases; HMT, histone methyltransferases; SUMO, small ubiquitin-like modifier; ATP, adenosine triphosphate.
suppressor genes through mutation or deletion. Epigenetic modifications, such as DNA methylation, serve a similar role to somatic mutations, leading to gene silencing. Overall, tumors tend to exhibit hypomethylation of repeat elements and pericentromeric regions and hypermethylation of normally unmethylated CpG islands. Gene-specific hypermethylation of CpG islands is associated with chromosomal instability, resulting in increased mutation rates and abnormal gene expression. For example, hypermethylation of critical genes, such as tumor suppressors, adhesion molecules, DNA repair genes, inhibitors of angiogenesis, and inhibitors of metastasis, is observed in tumorigenesis (7). Hypermethylation of DNA repair genes (e.g., hMLH1) is associated with genomic instability through disruption of chromosome replication control (7). Unlike gene mutations, DNA methylation can be reversed by DNMT inhibitors.

An important component of the cancer epigenome is an altered DNA methylation pattern of global demethylation and localized promoter hypermethylation. Aberrant methylation patterns are seen in almost every cancer type (7,16). In addition to aberrant methylation, tumors also show altered expression of HATs, making them another potential target for therapy (11). It is believed that histone modifications precede promoter hypermethylation to initiate transcriptional silencing of tumor suppressor genes. Repressive histone modifications are first established. Over time the density of DNA methylation increases in the gene promoter regions, which creates a “permanent” mode of silencing (17).

THE ROLE OF EPIGENETICS IN BREAST CANCER

In the United States, breast cancer is one of the most common cancers, with approximately 200,000 women diagnosed with breast cancer each year (http://www.cancer.org). In addition to well-understood DNA mutations in breast tumors, epigenetic alterations of gene expression are key contributors (16,18–24). A large number of genes are found to be hypermethylated in breast tumors but not in normal breast tissue. Many of these genes have well-characterized functions in tumor development. Genes commonly methylated in breast cancer and their potential roles in tumor development are listed in Table 2. Four key genes, RASSF1A, BRCA1, ESR1, and ESR2, are further discussed in detail.

RASSF1A

The CpG island of the Ras-associated domain family 1A gene (RASSF1) is hypermethylated in 60% to 77% of breast cancers, and its transcripts are frequently inactivated in cancer cell lines and primary tissues (13,22). When cells lacking RASSF1A expression are treated with a DNMT inhibitor, such as 5-aza-2’-deoxycytidine, expression can be reactivated (25). RASSF1A is involved in
growth regulation and apoptotic pathways. Specifically, RASSF1A is a Ras effector and induces apoptosis through its interactions with pro-apoptotic kinase MST1. Mouse knockout studies show that RASSF1a–/– mice are prone to spontaneous development of lung adenomas, lymphomas, and breast adenocarcinomas. Knockout RASSF1A mice are prone to early spontaneous tumorigenesis and show a severe tumor susceptibility phenotype compared to wild-type mice (25). In addition to breast tumors, hypermethylation of RASSF1A can be detected in nonmalignant breast cells and patient sera. In one study, RASSF1A hypermethylation in sera of breast cancer patients was detected in six out of 26 cases. These data suggest that RASSF1A hypermethylation may be a promising biomarker to screen putative cancer patients (25).

**BRCA1**

Germline mutations of BRCA1 and BRCA2 are responsible for familial breast cancers. Somatic mutations in sporadic cases are rare, but chromosomal losses occur in 30% to 50% of sporadic tumors, respectively (26). BRCA1 acts as a tumor suppressor gene for both breast and ovarian cancer (16). It encodes a multifunctional protein involved in DNA repair, cell cycle checkpoint control, protein ubiquitinylation, and chromatin remodeling (19). In vitro studies showed that decreased BRCA1 expression in cells led to increased levels of tumor growth, while increased expression of BRCA1 led to growth arrest and apoptosis. Inactivation of BRCA1 by promoter methylation is seen in 7% to 31% of

### Table 2  Genes Methylated in Breast Cancer

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cancer hallmark contribution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RASSF1A</td>
<td>Apoptosis</td>
<td>15</td>
</tr>
<tr>
<td>CYP26A1</td>
<td>Microenvironment, encodes for cytochrome p450 enzymes</td>
<td>17</td>
</tr>
<tr>
<td>KCNAB1</td>
<td>Microenvironment</td>
<td>17</td>
</tr>
<tr>
<td>SNCA</td>
<td>Microenvironment</td>
<td>17</td>
</tr>
<tr>
<td>RARβ</td>
<td>Growth regulation</td>
<td>15</td>
</tr>
<tr>
<td>TWIST</td>
<td>Cell lineage and differentiation</td>
<td>15</td>
</tr>
<tr>
<td>Cyclin D2</td>
<td>Cell cycle regulation</td>
<td>15</td>
</tr>
<tr>
<td>APC</td>
<td>Cell cycle regulation and apoptosis</td>
<td>8</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Cell cycle regulation</td>
<td>8,15</td>
</tr>
<tr>
<td>BRCA2</td>
<td>Tumor suppressor gene</td>
<td>15</td>
</tr>
<tr>
<td>p16</td>
<td>Cell cycle regulation</td>
<td>8,15</td>
</tr>
<tr>
<td>DAPK</td>
<td>Apoptosis</td>
<td>8</td>
</tr>
<tr>
<td>GSTP1</td>
<td>Cellular metabolism and carcinogen detoxification</td>
<td>8</td>
</tr>
<tr>
<td>TMS1</td>
<td>Apoptosis</td>
<td>8</td>
</tr>
<tr>
<td>PR</td>
<td>Growth regulation</td>
<td>8</td>
</tr>
<tr>
<td>CDH1</td>
<td>Angiogenesis</td>
<td>8</td>
</tr>
<tr>
<td>WWOX</td>
<td>Apoptosis</td>
<td>38</td>
</tr>
</tbody>
</table>
sporadic breast cancer cases (27). It is believed that aberrant methylation pairs with loss of heterozygosity to reduce BRCA1 expression in invasive sporadic breast tumors (16,19,22,26,28). The magnitude of the decrease of functional BRCA1 protein correlates with disease prognosis (19,22). Phenotypically, BRCA1-methylated tumors are similar to tumors with germline mutations. Higher-grade tumors tend to show complete loss of BRCA1 protein. In contrast to BRCA1, BRCA2 does not show as high a degree of promoter methylation (16).

**ESR1 and ESR2**

Along with other genes, the estrogen receptors (ERs) ERα and ERβ have been implicated in breast cancer development. ERα is encoded by ESR1, and when estrogen is activated, it stimulates cell proliferation. ERβ is encoded by ESR2 and is known to inhibit the proliferation and invasion of breast cancer cells. ERα and ERβ both have promoter-associated CpG islands that can be abnormally methylated in breast cancer, but ESR2 methylation has been less well studied (16,18). Almost all breast cancers show some degree of DNA methylation of the ESR1 gene promoter, but this methylation is only associated with gene repression in ~30% of tumors (16,18). Approximately 66% of breast cancers express ERα. A fraction of breast cancers that are initially ERα positive lose ER expression during tumor progression, but it is unclear if this is due to methylation or other causes (18).

The presence or absence of ERs plays an important role in both therapy and survival (16,18). ER-positive breast cancer can be treated with antiestrogenic drugs, such as tamoxifen, which binds to the ER, preventing estrogen from binding. This results in a decrease in growth of ER-positive breast cancer cells. ER-negative cells are no longer responsive to estrogen; therefore, antiestrogenic drugs have no effect (18). Reversal of ER-negative status is an attractive target for breast cancer therapy.

**BIOMARKERS AND DETECTION**

Gene-specific epigenetic changes for breast cancer are likely to occur early in tumorigenesis and have potential to be used as early detection and prevention methods (2). As epigenetic modifications become clear biomarkers for breast cancer treatment and therapeutic intervention, it is important to understand the many different techniques available for studying and detecting the presence of methylation, histone modifications, and microRNAs (11).

New, promising high-throughput methylation detection methods are available, which enable researchers and clinicians to identify an “epigenetic signature” specific to breast cancer. Biomarkers based on unique DNA methylation and histone modification patterns of breast tumors may be used in the future for diagnosis and early detection. Currently, breast cancer detection relies on various screening methods. Cytology is the main standard for identification
of abnormalities typical of cellular transformation. Quantitative multiplex methylation-specific polymerase chain reaction (PCR) (QM-MSP) is a highly sensitive method to quantitate cumulative gene promoter hypermethylation in samples where DNA is limiting. First, gene-specific primer pairs are used to coamplify genes, independent of their methylation status. Next, gene-specific primers (methylated and unmethylated) conjugated with fluorescent labels are combined with the first PCR product. The signal generated is proportional to the extent of DNA amplification (15). Since it is known that multigene methylation of CpG islands is common in early breast cancer, a panel of nine genes were tested for their sensitivity and specificity of detecting premalignant changes using QM-MSP. These genes included RASSF1A, RARβ, TWIST, HIN1, Cyclin D2, APC, BRCA1, BRCA2, and p16. Tests were done using ductal lavage, nipple aspiration fluids, and fine needle aspirates.

Using this detection panel, the rate of detection of breast cancer cells rises from 43% sensitivity with cytologic examination alone to 71.4% sensitivity (15). Hypermethylation of genes that are commonly methylated in breast cancer in sera in breast cancer patients has also been used to detect early malignant changes. Sera methylation can be detected using methylation-specific PCR (MS-PCR), a noninvasive PCR-based technique. Early studies indicate that MS-PCR is a promising approach to screen putative cancer patients (25). This technique utilizes primers designed for methylated or unmethylated bisulfite-modified DNA (29). RASSF1A gene hypermethylation has been detected using MS-PCR in sera of ovarian cancer patients, with 100% specificity (25).

Recurrence rates for breast cancer in the same breast have been recorded as high as 40% despite negative pathological margins. Researchers believe that a primary tumor may serve as a locus from which methylation density progressively diffuses outward to surrounding tissues. CYP26A1, KCNAB1, RASSF1A, and SNCA have found to be frequently methylated in adjacent normal tissue. A methylation profile containing genes such as these, or those identified in other studies (Table 2), could be used by clinicians to determine whether patients should undergo radiological therapy or other treatments to prevent future recurrence (17).

MassARRAY and pyrosequencing, two additional methods of detecting methylation, are currently being developed for use in clinical settings. MassARRAY is a highly sensitive technique that uses base-specific cleavage and matrix-assisted laser desorption/ionization time-to-flight mass spectrometry (MALDI-TOF MS). It is capable of detecting DNA methylation levels as low as 5%. It is suitable for testing methylation patterns on various sources, including archival tissues and laser capture microdissected specimens (1). Pyrosequencing is a locus-specific quantitative method that utilizes the detection of pyrophosphate, which is liberated from incorporated nucleotides by DNA polymerase during strand elongation. Free pyrophosphates are converted to adenosine triphosphate (ATP), which provides energy for the oxidation of luciferin to then generate light. Nucleotides are added sequentially to enable base calling.
Pyrosequencing has two major advantages over MS-PCR. First, the data are actual nucleotide sequences rather than fluorescence data. Second, pyrosequencing can detect partially methylated sequences that are outside of the priming sites (30).

Commonly used methods for detecting histone modifications usually rely on chromatin immunoprecipitation (ChIP). In ChIP, DNA is cross-linked to chromatin, followed by co-immunoprecipitation using antibodies against specific histone residues or chromatin proteins. “ChIP” DNA can be used in microarrays, cloning, and sequencing (11). Techniques for detecting histone modifications have not yet been adapted for clinical application.

**CLINICAL APPLICATION**

Breast cancer prevention, treatments, and diagnostics are being developed to target epigenetic changes leading to breast cancer. Treatments for breast cancer currently being studied extensively focus on reversing aberrant DNA methylation and histone acetylation of tumor suppressor genes and genes involved in therapeutic response. Combinations of epigenetic-targeted therapies with conventional chemotherapeutic agents may provide a way to resensitize drug-resistant tumors (Table 3).

**Methylation Therapy**

DNA methylation in proliferating cells is dependent on continued expression of DNMTs. Treatments that focus on inhibiting the expression of these proteins would therefore result in progressive reduction in DNA methylation in newly divided cells. There are three main classes of DNA methylation inhibitors: nucleoside inhibitors, nonnucleoside weak inhibitors, and rationally designed inhibitors. Nucleoside inhibitors trap DNMTs for degradation. However, they require DNA incorporation and active DNA synthesis, which limits the drug in hypoproliferating cells. 5-Azacytidine (5-aza-dC) and its deoxy analog are two of the first hypomethylating agents approved by the U.S. Food and Drug Administration for cancer treatment (31). They work by two mechanisms: (i) as demethylators by incorporating into CpG sites opposite a methylated CpG site on the template strand and (ii) by binding to the cytosine in the catalytic domain of the DNMT1 enzyme. Within a few hours of treatment with 5-aza-dC, there is rapid passive loss of methylation (32). Nonnucleoside inhibitors are not as well understood and are limited by their low level of hypomethylation induction. Nonnucleoside inhibitors were first discovered for their use in reducing auto-reactivity for autoimmune diseases. During this study researchers discovered their potential as demethylating agents. Researchers are currently using clinically approved nonnucleoside inhibitors as a starting point to study their use as demethylating agents. Rationally designed inhibitors are also in development. These drugs are designed on the basis of knowledge of specific chemical
### Table 3  Drugs in Use for Breast Cancer Treatment

**Nucleoside inhibitors**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Use</th>
<th>Research level</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-Azacytidine</td>
<td>Trap DNMTs for degradation</td>
<td>Clinical use</td>
<td>31,40</td>
</tr>
<tr>
<td>5-aza-2’dexoxytidine</td>
<td>Trap DNMTs for degradation</td>
<td>Clinical use</td>
<td>31</td>
</tr>
<tr>
<td>Zebularine</td>
<td>Unknown mechanism</td>
<td>Shown promise in vitro</td>
<td>31</td>
</tr>
<tr>
<td>5′-fluoro-2’dexoxytidine</td>
<td>Unknown mechanism</td>
<td>In clinical trials</td>
<td>31</td>
</tr>
</tbody>
</table>

**Nonnucleoside inhibitors**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Use</th>
<th>Research level</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procainamide</td>
<td>Mechanism not well understood but proven to decrease expression of DNMT1 and DNMT3a</td>
<td>Clinically used as an antihypertensive drug</td>
<td>31,41,42</td>
</tr>
<tr>
<td>Hydralazine</td>
<td>Mechanism not well understood but proven to decrease expression of DNMT1 and DNMT3a</td>
<td>Clinically used as a vasodilator</td>
<td>31,41,42</td>
</tr>
</tbody>
</table>

**Rationally designed inhibitors**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Use</th>
<th>Research level</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSC303530</td>
<td>Blocks the active site of DNMTs</td>
<td>Basic research</td>
<td>31,43</td>
</tr>
<tr>
<td>NSC401077</td>
<td>Blocks the active site of DNMTs</td>
<td>Basic research</td>
<td>31,43</td>
</tr>
</tbody>
</table>

**HDAC inhibitors**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Use</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Research level</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAQ824</td>
<td>HDAC inhibitor</td>
<td>Unknown</td>
<td>Important to pay attention to schedule of administration</td>
<td>Preclinical level; studies on hematological malignancies</td>
<td>40</td>
</tr>
<tr>
<td>Magnesium Valporate and Hydralazine</td>
<td>HDAC and methyltransferase inhibitor</td>
<td>Increases efficacy of conventional cytotoxic agents</td>
<td>Unknown</td>
<td>Phase III clinical trials</td>
<td>44</td>
</tr>
<tr>
<td>N-hydroxy-2-propenamide</td>
<td>HDAC inhibitor</td>
<td>Suppresses mammary tumor growth in cells</td>
<td>Unknown</td>
<td>Research level</td>
<td>145</td>
</tr>
</tbody>
</table>

*Abbreviations: DNMTs, DNA methyltransferases; HDAC, histone deacetylase.*
responses in the body, tailoring combinations of these to fit a treatment profile. An example of a rational design involves the use of three-dimensional information to target the drug. Limitations of DNA methylation inhibitors include the cooperation between different DNMTs such that the drugs must inhibit several of them simultaneously. All demethylating agents suffer from nonspecificity and toxicity. The next phase of drug development is how to target gene-specific hypomethylation (e.g., unmethylated oligonucleotides) (31).

**Histone Modification Therapy**

Histone deacetylases (HDACs) are also targets for therapy (Fig. 3). HDACs are important because of their role in cell cycle arrest and inhibition of apoptosis. These drugs have been studied in isolation and synergistically with DNMTs. Therapies that target both DNMTs and HDACs act globally, which can cause reactivation of other genes that are typically silenced, such as oncogenes (18,33–37).

**ER Therapy**

Drugs are in development for different types of breast cancer. A therapy that could restore gene expression of *ERS1* (for ER-negative cancer patients) could reestablish cancer cell growth regulation through estrogen. After reexpression of *ERS1*, antiestrogenic drugs could subsequently be used. A number of drugs, including DNMT inhibitors and HDAC inhibitors, are being used to reactivate *ER* expression (Table 4). DNMTs and HDACs work synergistically to silence gene expression of *ER*. One problem found in treatments that target DNMTs and HDACs is that alone they may not be enough to reverse *ER*-promoter hypermethylation.
Tamoxifen is a commonly used drug for ERα-positive breast cancers. Unfortunately, acquired resistance of tamoxifen occurs in nearly 40% of patients. Therefore, it is important to identify patients most likely to respond to Tamoxifen and to identify individuals likely to acquire resistance. One promising biomarker for Tamoxifen response is WWOX. Preliminary studies show that WWOX expression levels predict tamoxifen resistance better than the two previously known biomarkers, PR and HER2 (38). WWOX appears to mediate tamoxifen sensitivity, and its expression is reduced in a large fraction (63.2%) of breast cancers (39). The primary mechanism of downregulation of WWOX is through DNA hypermethylation of its regulatory region.

### FUTURE DIRECTIONS

The next 10 years are likely to bring significant advances for breast cancer treatment. Therapeutics that target methylation and histone modifications may play a leading role in treating breast cancer. Since epigenetic modifications can also be used as biomarkers, targeted therapies may some day be used as preventative measures.

### REFERENCES


### Table 4  Drugs Involved in Reactivation of ER Expression

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Result</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-aza-Dc (methyltransferase inhibitor)</td>
<td>Partial demethylation of ER CpG islands, reexpression of ER mRNA, and synthesis of functional ER protein</td>
<td>18,36,37,46</td>
</tr>
<tr>
<td>Antisense oligonucleotides (for DNMT1)</td>
<td>ER gene reexpression and estrogen responsiveness reestablished</td>
<td>18,35</td>
</tr>
<tr>
<td>TSA (HDAC inhibitor)</td>
<td>Reexpression of ERα protein</td>
<td>18,33</td>
</tr>
<tr>
<td>DNMT and HDAC inhibitor</td>
<td>Reactivation ERα protein and expression</td>
<td>18,37</td>
</tr>
<tr>
<td>Scriptaid (HDAC inhibitor)</td>
<td>Reexpression of functional ER in vivo</td>
<td>18,34</td>
</tr>
<tr>
<td>Scriptaid and 5-aza-Dc</td>
<td>More effective in inducing ER expression than just Scriptaid or 5-aza-Dc alone</td>
<td>18,34</td>
</tr>
</tbody>
</table>

Abbreviations: ER, estrogen receptor; DNMT, DNA methyltransferase; HDAC, histone deacetylase; TSA, trichostatin A.
The Role of Epigenetics in Breast Cancer


INTRODUCTION

The progression of normal breast epithelial cells from a normal state toward one characterized by uncontrolled growth and metastatic behavior is caused by the deregulation of key cellular processes and signaling pathways. These alterations in normal cellular behavior are rooted in the accumulation of genomic and epigenomic lesions that impact hallmarks of cancer, such as the ability of the cell to control proliferation, undergo apoptosis, increase motility leading to invasion, and alter angiogenesis. A suite of technologies has now been developed to assess genomic and epigenomic aberrations that contribute to cancer progression. The application of these has shown that genome copy number abnormalities (CNAs) are among the most frequent genomic aberrations. Remarkably, these studies have revealed that 10% to 15% of the genes in a typical carcinoma tumor may be deregulated by recurrent genome CNAs and regions. Some of these genes influence disease progress and so may be assessed to facilitate prognosis. Others influence response to therapy and so may be assessed as predictive markers. Some of these enable oncogenic processes, on which tumors depend for survival, and so are candidate therapeutic targets. In most cases, array comparative
genomic hybridization (CGH) is the method of choice for their discovery and may be used in some setting for clinical assessments of these abnormalities. Accordingly, we review here several of the current array CGH technologies available and the considerations needed when determining the technology most applicable to a given study.

CGH analyses of thousands of primary breast tumors and cell lines have now defined recurrent CNAs throughout the genome and identified genes encoded in these aberrant regions that are deregulated by the abnormalities (1,2–4). In breast cancer, for example, 10% to 15% of known genes are deregulated by these CNAs (2–4). The aberrations tend to fall into two general classes that may contribute in different ways to cancer pathophysiology: low-level CNAs that occur through the gain or loss of extended regions of single chromosomes or extended regions thereof and high-level amplification or homozygous deletions that tend to be narrower in genomic extent.

Most genes are deregulated by low-level aberrations. The contributions of these genes to cancer pathophysiology are not well understood; however, evidence is now emerging that these may provide, en ensemble, a general metabolic advantage (4–6). These aberrations typically occur relatively early in the carcinogenesis process. It has been proposed that these aberrations are selected during successful transition through a telomere crisis (7).

High-level amplification and narrow homozygous deletions provide strong evidence of active selection during cancer progression and seem to occur later during cancer progression. In breast cancer, regions of high-level amplification encode well-studied oncogenes, such as ERBB2 (17q12) (8), FGFR1 (8p11) (9), PVT1 and MYC (8q24) (10,11), CCND1 and EMS1 (11q13) (12), and ABI1 and ZNF217 (20q13) (13,14), while deletions contribute to the inactivation of tumor suppressor genes such as TP53 (17p13) (15), CDKN2A (9p21) (16), and BRCA1 (13q14) (17). While it is often assumed that these oncogenes and tumor suppressor genes are the sole targets for these aberrations, a growing body of functional data suggests that several genes in each region of recurrent abnormality may be coselected by the event. In breast cancer, for example, regions of amplification in which more than one gene appears to be functionally important include 8p11, 8q24, 11q13, 17q21, and 20q13 (4). Table 1 summarizes published studies for genes in these regions of amplification. The concept of multigene selection is also supported by studies in mouse models, where the extents of amplicons in mice and humans are similar. Similarities in the extent of amplification around ERBB2 in humans and in polyomavirus middle T-induced mammary tumors in mice are particularly striking (18). Similar studies are emerging at other organ sites (19). As a result, it is critically important that the extent of each region be defined as precisely as possible. High-resolution CGH analyses are now contributing to this process.

Several recurrent high-level CNAs are associated with clinical endpoints, including reduced survival duration and response to aberration-targeted therapies. Assays for recurrent aberrations or genes deregulated by CNA abnormalities have
already been incorporated into molecular indicators of outcome. In addition, several therapeutic agents are now being developed to attack genes deregulated fully or in part by recurrent genome abnormalities so that assays for these targets may be useful as predictors of response. In breast cancer, the most prominent examples are trastuzumab (20,21) and lapatinib (22) (NEJM PMID 17192538), developed to attack tumors driven by amplification of ERBB2. In addition, therapies have been developed to treat tumors arising from deletion or inactivation of TP53 (23) and activation of elements of the PI3-kinase pathway, including PIK3CA, PTEN, AKT, and ZNF217 (24–27).

The ability to detect and define the extents of CNAs has steadily improved with the evolution of CGH analysis technologies (Table 2). The first CGH studies mapped CNAs onto metaphase chromosomes, but the resolution of such analyses was limited to about 10 Mbp by the complex packaging of the DNA into chromosomes (28). The introduction of array CGH in which CNAs were mapped onto arrays of cloned probes substantially increased resolution. Early arrays comprised hundreds to thousands of probes and improved resolution to about 1 Mbp (29,30). Current cloned probe arrays interrogate the genome at approximately single-gene resolution (5). Today, several commercial platforms map abnormalities onto arrays of oligonucleotide probes with subgene resolution (either on flat substrates or attached to microbeads) (31), and some allow analyses in an allele-specific manner (32). Applications of these platforms are further refining the extents of recurrent genomic aberrations but also reveal the somewhat surprising existence of germline copy number polymorphisms that must be taken into account when interpreting CGH measurements. Several aspects of these important array CGH analysis technologies are reviewed in the following sections.
COMPARATIVE GENOMIC HYBRIDIZATION

CGH is the primary technology used for assessment of genome copy number. Typically, analysis proceeds by hybridizing differentially labeled tumor and normal DNA to a representation of the normal genome, although some platforms perform tumor and normal reference hybridization in separate representations. The ratio of tumor label to normal label, hybridized to each locus on the representation is then measured as an estimate of relative copy number at that locus.

The first CGH experiments employed metaphase chromosomes as the representation onto which information was mapped (28,33). While useful, the genomic resolution was limited by the nonlinear organization of DNA along metaphase chromosomes. Thus, chromosome CGH has now been largely supplanted by array CGH, in which information is mapped onto arrays of DNA probes (Fig. 1). The initial array CGH platforms comprised large genomic cloned probes, such as yeast artificial chromosomes (YACs) (34), bacterial artificial chromosomes (BACs), and P1-derived artificial chromosomes (PACs) (29), or shorter cDNA sequences spotted on glass slides (35). More recently, synthetic oligonucleotide probes have been used successfully (36–39). Several platforms are summarized in Table 1 and in the following sections.

Cloned probe arrays are typically available from large academic cores with the experience and resources to maintain and print the arrays. These arrays are developed using the large collections of cDNAs (35) or BACs (29,30,40) developed during the course of the human genome program. The advantages of

<table>
<thead>
<tr>
<th>Platform</th>
<th>Program</th>
<th>Reference</th>
<th>Type</th>
<th>Required software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-color arrays</td>
<td>aCGH</td>
<td>54</td>
<td>HMM</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>GLAD</td>
<td>56</td>
<td>Adaptive weight</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>CLAC</td>
<td>57</td>
<td>Clustering along</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>DNAcopy</td>
<td>55</td>
<td>Segmentation</td>
<td>R</td>
</tr>
<tr>
<td>Single-color arrays</td>
<td>dChipSNP</td>
<td>37</td>
<td>HMM</td>
<td>none</td>
</tr>
<tr>
<td>(Affymetrix)</td>
<td>CARAT</td>
<td>69</td>
<td>Regression</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>CNAG</td>
<td>70</td>
<td>Regression</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>GEMCA</td>
<td>71</td>
<td>Regression</td>
<td>none</td>
</tr>
</tbody>
</table>

Abbreviations: aCGH, array comparative genomic hybridization; HMM, hidden Markov model; GLAD, Gaussian model–based approach; CART, copy number analysis with regression and tree; CLAC, clustering along each chromosome; SNP, single nucleotide polymorphism; CNAG, Copy Number Analyzer for GeneChip; GEMCA, Genotyping microarray based copy number variation analysis.
this approach are that the clones are readily available in the public domain and the signals produced during hybridization are large so that measurement precision is relatively high. The disadvantages are that the clones are not perfectly curated and mapped, so some may be mismapped. In addition, managing the large number is daunting, and the ultimate resolution of the analysis platforms is limited by the relatively genomic extents of the cloned probes.

Commercial vendors, such as Affymetrix, Agilent, Illumina, and NimbleGen, have begun offering oligonucleotide-based arrays with which to study genomewide CNAs (41–43). The Illumina arrays are unique in that the oligonucleotide probes are attached to microbeads (39), while the NimbleGen arrays are synthesized on demand and so can be easily customized (44). The smaller target sequences, 25 to 85 bp, also offer the possibility of higher resolution than is possible using cloned probe arrays. Some protocols generate relative copy number information by comparing tumor and normal hybridization signals using two-color cohybridization approaches, while others rely on hybridizations of tumor and normal reference DNA to separate arrays. The latter approach is technically simpler but requires additional normalization steps to account for array-to-array differences, introducing additional experimental variables when analyzing data sets.
Since each oligonucleotide probe is synthesized to be complementary to a known sequence, oligonucleotide array platforms do not suffer as much from incorrect mapping, and probes can be selected to be uniquely represented in the genome, thereby minimizing cross-hybridization artifacts. The smaller sequences also are useful in detecting single nucleotide polymorphisms (SNPs), since hybridization rates to 25 mer are significantly reduced by a single nucleotide mismatch (37,45). The additional information provides researchers the ability to detect loss of heterozygosity (LOH) in regions of neutral copy change and also provide data for use in association studies (42). Regions of LOH that do not carry imbalances in copy number may provide evidence of selective pressures duplicating one allele coupled with the loss of the other. This is an important mechanism enabling removal of wild-type alleles while selecting mutations. Allele-specific selection can be either positive or negative. For example, loss of the wild-type allele and duplication of an inactivating mutation might lead to inactivation of a tumor suppressor gene. The high frequency of LOH on 17p where p53 is located is an example of this type of selection (Fig. 2). On the other hand, amplification of a polymorphism that increases gene activity may contribute an advantage to the tumor. Amplification of a specific polymorphism in the aurora kinase STK15 is one example (46). The utility of allele-specific copy number analysis may also contribute to the identification of polymorphisms that confer increased susceptibility to cancer, since evidence is now emerging that tumors may increase the number of copies of polymorphisms that confer increased risk and may decrease the number of copies of polymorphisms that confer decreased risk of developing cancer (47). The frequency of allele-specific copy increases and decreases in individual tumors is substantial, as illustrated in Figure 3.

While commercial oligonucleotide arrays offer advantages, the smaller target sequences also present some disadvantages. Foremost among these is the relatively low signal-to-noise ratio generated during hybridization. Because the hybridization sequences are smaller, there is less nucleic acid sequence available to label, leading to less intense labeling. This typically leads to higher variability between probe sets in oligonucleotide arrays, and the need to combine multiple data points to estimate copy number, which in turn reduces the actual resolution of these arrays (38,48). The increased variability compared to BAC arrays also makes the detection of low-level changes, such as single-copy gains and losses, more difficult to detect, although performance with oligonucleotide arrays is improving steadily. Typically, microgram quantities of starting DNA are needed for an analysis.

The most recent array CGH analysis platform does not operate through hybridization of tumor and normal DNA to a normal representation. Instead, it utilizes a molecular inversion probe (MIP) interrogation strategy to perform allele-specific copy number analysis (45). In this approach, each locus is interrogated by a linear MIP probe that has sequence homology to approximately 20-bp regions flanking a single base to be tested. Each MIP also contains a unique
tag sequence and universal polymerase chain reaction (PCR) primer-binding sites facing away from each other so that PCR with primers to the universal sequences will not amplify. The first step in MIP CGH is to hybridize the linear MIPs to the DNA sample to be interrogated. After hybridization, the reaction mix is separated into multiple tubes, each containing polymerase, ligase, and one of the four nucleotides. The polymerase and ligase then fill the gaps in each reaction mix that contains a nucleotide complementary to the base at the test locus. The linear DNA and probes are then eliminated, reactions are mixed, and the ligated probes are amplified and labeled using PCR with the universal primers. The resulting PCR products are hybridized to an array comprising oligonucleotides homologous to the tag sequences in the MIPs. To date, MIP analyses interrogating over 20,000 loci have been reported (45). This approach reports
Figure 3 Allelic copy number changes in breast cancer cell lines from a 50K MIP array platform. Allele copy numbers are represented as the higher copy allele and lower copy allele. (A) Allele copy number data from HCC38 from chromosome 5. The lack of copies of the lower allele is indicative of LOH. The array detects a region of LOH without copy number change. (B) Allele copy number data from ZR75.30 from chromosome 8. 8q displays a region of monoallelic amplification in the region that contains MYC. Abbreviations: MIP, molecular inversion probe; LOH, loss of heterozygosity.
copy number in an allele-specific manner and works well with FFPE material since only ~40 bp of intact DNA is needed at each locus for proper analysis. Typically, only about 75 ng of DNA is needed for an analysis since the reaction itself contains a PCR amplification step. The approach has the advantage that the tag sequences and arrays can be designed to have uniform hybridization conditions to copy number and the hybridization does not depend on the quality of the hybridized DNA.

DNA REQUIREMENTS

One of the most important factors when considering a clinical CGH tool is the quality and amount of DNA available for analysis. Depending on which array platform is chosen, the amounts needed for successful hybridization can range from 50 ng to 3 μg. In addition, some protocols have been developed that allow analysis of material from formalin-fixed, paraffin-embedded (FFPE) samples (39,49). The largest amounts of starting material generally are used in protocols in which entire genomes are labeled and hybridized to the array, as is usual for array CGH using BAC or Agilent oligonucleotide arrays. In these experiments, the starting DNA is typically labeled with fluorescent nucleotides using random priming. For BAC arrays, unlabeled repeat-rich DNA is included to suppress hybridization of repeated sequences to repetitive sequences that are present in the BAC probes. Larger amounts of DNA are needed to “drive” these whole-genome hybridization reactions. The amount of starting material needed for analysis is reduced if the DNA is amplified or reduced in complexity—for example, through a PCR amplification step preceded by restriction digestion that reduces the complexity of the hybridized DNA (50). This complexity reduction results in an increase of template sequences, which can utilized in downstream hybridization steps. Both academically and commercially developed platforms result in data sets that are able to detect gains and losses with the caveat that variations and noise may differ between platforms (Fig. 4).

While there is a range of starting DNA amounts between the various platforms, recent protocols have been developed that can amplify a large quantity of DNA from small starting amounts. These applications can be critical when dealing with samples that are limiting. These whole-genome amplifications are useful for both intact, high-quality DNA and degraded samples and involve the use of φ29 enzyme or PCR (51,52). Alternately, the MIP platform has a targeted PCR amplification inherent to the protocol, so it is particularly tolerant of small amounts of starting DNA (45). To date, successful data sets have been generated using as little material as 1 ng or a single cell (53). However, as with any random amplification procedure, the protocols may artificially create misrepresentations of the genome, resulting in false positives for amplifications and deletions. Determining the amount and quality of starting material along with the
Figure 4 Chromosomal copy number changes in the breast cancer cell line MDAMB453. DNA from MDAMB453 is applied to three separate platforms: (A) a 3K BAC array, (B) a 20K MIP array, and (C) a 50K XbaI Affymetrix SNP array. Data from each platform are scaled using adjacent data points so that the resulting resolutions are comparable. The data are then smoothed using the GLAD software package. Abbreviations: BAC, bacterial artificial chromosome; MIP, molecular inversion probe; SNP, single nucleotide polymorphism; GLAD, Gaussian model–based approach.
Comparative Genomic Hybridization in Breast Cancer

type of information required will help determine the appropriate array CGH platform and DNA preparation technique to utilize.

ABNORMALITY DETECTION AND INTERPRETATION

Computational strategies to identify copy number gains and losses are a critical part of array CGH. The first step in the analysis of data is the identification of any features with anomalous intensity readings. These could result from a high background, errors in printing, contamination, or air bubbles prohibiting hybridization of the sample to the array. After removal of these spurious data points, a normalization step is applied to calculate the tumor hybridization intensity along the genome, relative to the normal hybridization intensity. A statistical analysis is then applied to identify genome segments with different copy number levels. Several strategies have been developed for this purpose and are summarized in Table 2. These include the hidden Markov model (HMM) (54), circular binary segmentation (55), a Gaussian model–based approach (GLAD) (56), hierarchical clustering along each chromosome (CLAC) (57), an expectation maximization–based method (58), a Bayesian model (59), and a wavelet approach (60). Several of these algorithms have been made available as R packages through Bioconductor (www.bioconductor.org). Comparisons of several of these methods have now been published (61,62).

Detection of the differences in copy number between tumor and normal samples is increasingly straightforward and precise as CGH procedures improve. However, the interpretation of CGH analyses requires additional interpretation. First, high-resolution array CGH analyses are important now that focal, germline genome copy number polymorphisms are relatively common (36). These may be scored as CNAs if the samples for tumor versus normal CGH analyses are not from the same individual. In addition, it is tempting to think that the regions of the genome not detected as aberrant during CGH analyses are present in cancer genomes in two copies. This is often not the case since tumor genomes may be highly aneuploid. Thus, a CGH comparison of a perfectly tetraploid tumor with a diploid normal sample will not show any evidence of abnormality. Thus, if absolute copy number assessments are important, it will be necessary to normalize the CGH analyses using information about absolute copy number at one or a few loci measured using fluorescent in situ hybridization (FISH) or quantitative PCR.

CONCLUSIONS AND FUTURE DIRECTIONS

Array CGH technologies are becoming increasingly powerful tools for the identification of recurrent genome CNAs that are associated with outcome or response to therapy in breast and other cancers (63). The ability to detect genetic copy number changes of both large and small scales will be critical in evaluating the most effective therapies on an individual patient basis. In the short term, array CGH will likely be accomplished using commercial oligonucleotide analysis platforms that allow assessment of allele-specific copy number using
material from FFPE samples. In the long term, however, array CGH may be supplanted by high-throughput sequencing technologies that allow assessment of both copy number and DNA sequence through deep sequencing. It is already clear that shotgun sequencing, if done at sufficient redundancy, offers the same or better resolution than CGH platforms (64,65). Next-generation sequencing platforms based on massively parallel single-molecule sequencing promise sufficiently low-cost sequencing, so that these platforms may become a viable alternative to array CGH (66,67). Today, this is not cost effective, but costs of high-throughput sequencing are falling rapidly, so these approaches should be followed closely.

REFERENCES


INTRODUCTION

The selection of adjuvant chemotherapy for breast cancer is currently more of an art than a science. Several combination chemotherapy regimens can improve disease-free and overall survival of stage I to stage III breast cancer when used as adjuvant (postoperative) or neoadjuvant (preoperative) therapy (1,2). However, no chemotherapy regimen is universally effective for all patients, and it is currently impossible to predict which of these regimens will actually work for a particular individual. Choosing the best possible adjuvant or neoadjuvant chemotherapy regimen for an individual is important because the right choice may make the difference between cure and relapse. Results from traditionally designed randomized clinical trials offer little guidance for individualized treatment selection. These studies compare the activity of different treatments to establish which is the most effective for the entire study population. This approach ignores the fact that many patients are cured with earlier generation adjuvant regimens and therefore are not well served by the administration of
longer, potentially more toxic, and more expensive second- or third-generation combination therapies. These studies also cannot reveal whether some cancers are particularly sensitive to a specific drug (or a combination of drugs) and therefore can only be cured by that treatment. The most effective treatment for an individual may or may not be the same regimen that is the most effective for the whole population. These theoretical concerns and the increasing number of adjuvant chemotherapy options, as well as the advent of new molecular analytical techniques, fuel a renewed interest in developing chemotherapy response predictors that may lead to more personalized treatment recommendations.

THE ROLE OF NEOADJUVANT CHEMOTHERAPY IN THE TREATMENT OF BREAST CANCER

Administration of chemotherapy before surgical resection of the cancer (preoperative or neoadjuvant chemotherapy) provides a unique opportunity to identify molecular markers of response. Preoperative chemotherapy is the current standard of care for locally advanced breast cancer because this treatment often renders previously inoperable cancers amenable to surgery due to frequent and substantial tumor responses (3). Preoperative chemotherapy is also increasingly used in the management of operable breast cancer because it can reduce tumor size and allow for more limited surgery and possibly improved cosmetic outcome (4). Importantly, there is no statistically significant difference in survival between those who receive chemotherapy before surgery (i.e., neoadjuvant) and those who are treated with the same regimen postoperatively (i.e., adjuvant) (5–7). Therefore, preoperative chemotherapy is an appropriate option for most patients who present with a breast cancer that requires systemic therapy. Furthermore, eradication of all invasive cancer from the breast and regional lymph nodes by the preoperative therapy [i.e., pathologic complete response, (pCR)], is associated with excellent long-term, disease-free, and overall survival (7,8). Therefore, one could propose that the regimen most likely to induce pCR is the most effective treatment for a particular individual, and that such treatment could be ideally suited to that patient either as adjuvant or neoadjuvant therapy. This is the rationale why pCR is a commonly used clinical endpoint for breast cancer response predictors. However, it is important to realize that most patients with less than pCR had at least some clinical response to chemotherapy, and therefore cannot be considered truly resistant to treatment, but rather, cases with pCR represent extreme chemotherapy sensitivity.

CURRENT CHEMOTHERAPY RESPONSE PREDICTORS

There are several clinical and pathologic variables of breast cancer that are consistently associated with a greater likelihood of response to neoadjuvant chemotherapy. These include estrogen receptor (ER)-negative status, smaller tumor size, increased proliferative activity, and high nuclear or histological grade (9–12).
Tumor size, grade, and ER-status are routinely available parameters that can be combined into a multivariate response prediction model (13). One such model is freely available at our institutional Web site, http://www.mdanderson.org/Care_Centers/BreastCenter/dIndex.cfm?pn¼4448442B2-3EA5-4BAC-98310076A9553E63. However, clinical variable–based predictions are limited to predict response to chemotherapy in general and lack regimen specificity.

A large number of molecular markers were proposed as potential predictors of response to various chemotherapy agents (11,12,14–16). However, currently there are no molecular markers of response that could be used in the clinic to select patients for a particular therapy. Biomarker studies were traditionally conducted as correlative science projects attached to therapeutic trials as optional components. This has imposed serious limitations on these studies. Sample sizes for marker discovery were limited by the availability of specimens that has lead to under-powered observations and introduced potential bias due to subset analysis. Most commonly, only associations between marker and response were reported, but no formal prediction models were developed. The methodologies to detect a marker were rarely standardized, and definitions for test-positive and test-negative cases often differed from study to study, even for the same marker.

In addition to these technical limitations, there are important theoretical considerations, which also suggest that single gene molecular markers expressed by the cancer may have limited predictive value. Factors that reside outside the neoplastic cells could influence drug response. The importance of interactions between stroma and neoplastic cells, including neovascularization, changes in interstitial fluid pressure, and immune/inflammatory response are increasingly recognized as potential determinants of prognosis and possibly response to therapy (17,18). Rates of drug metabolism are also variable and can have an effect on drug activity. Cellular response to toxic insults is a complex molecular process that involves simultaneous induction of several proapoptotic and survival pathways, including activation of repair and drug elimination mechanisms (19,20). Tumor response may depend on the balance of these pathways, and therefore any single gene may contain only limited information about the complexity of the mechanisms that determine response to chemotherapy.

**GENE EXPRESSION PROFILING AS A NOVEL TOOL FOR RESPONSE PREDICTION**

Gene expression profiling with DNA microarrays represents a relatively recent tissue analysis tool that was developed in the early 1990s (21). This technology enables investigators to measure, in a semiquantitative manner in a single experiment, the expression of almost all mRNAs expressed in a small piece of cancer tissue. The rationale behind pharmacogenomic research is that by using

Pharmacogenomics is defined in this review as the use of mRNA expression profiles of the tumor to predict drug efficacy. Pharmacogenetics is defined as the study of DNA variations in the germline to predict toxicity or treatment response.
this technology, one can identify individually predictive genes and combine these into a multivariate prediction model that will be more accurate than any single gene alone. Multivariate outcome prediction models are not new to medicine. Independent and individual weakly predictive clinical variables have been combined into useful and validated prognostic or predictive models. There are several established statistical methods for combining variables into prediction models, including the commonly used logistic regression analysis and the Cox proportional hazards regression model (22). Predictors could be binary—that is, predicting response versus no response—or could predict the probability of response on a continuous scale. The Adjuvant Online (http://www.adjuvantonline.com) software represents a freely available validated multivariate prognostic prediction tool for breast cancer.

The general strategy for multigene marker development in neoadjuvant studies is to compare the transcriptional profiles of cancers that responded to therapy with those that did not and to identify differentially expressed genes that can be combined into a prediction model. There are several caveats that make pharmacogenomic predictor discovery challenging.

1. The multiple comparison problem inherent to microarray analysis is well known. It refers to the large number of variables (i.e., genes) that are compared between two small data sets. As a consequence, observation of many small $p$ values may be due to chance. This could result in spuriously identifying genes as associated with response when in fact they are not. Several methods used to adjust for this phenomenon, at least to some extent, by calculating false discovery rates for particular $p$ values were described (23).

2. Another confounder relates to the coordinated expression of thousands of genes. Individual transcripts do not represent independent variables, but rather their expression is highly correlated with one another. Many phenotypic characteristics of breast cancer are also associated with, and likely caused by, the coordinated expression of thousands of genes. For example, ER-positive cancers differ from ER-negative tumors in the expression of thousands of genes (24). The gene expression pattern of high-grade tumors is also different from low-grade tumors (25). These large-scale gene expression patterns associated with clinical phenotypes have a profound influence on the pharmacogenomic discovery process, because ER-negative, high-grade breast cancers are more sensitive to many different types of chemotherapies compared with ER-positive and low-grade tumors. Unadjusted pharmacogenomic comparison of responders with nonresponders will yield gene lists that are dominated by ER- and grade-associated genes, and the resulting pharmacogenomic predictor may represent a predictor of clinical phenotype. When the discovery sample size is large, it is possible to adjust for these phenotypic confounders; however, in small studies, these adjustments may be difficult to do.

3. The probability that a supervised pharmacogenomic discovery approach, when responders are compared with nonresponders, will lead to regimen-specific
predictors depends on to what extent the response groups are balanced for strong phenotypic markers and on the extent of molecular differences that determine drug-specific response. If drug sensitivity is influenced by the two- to threefold higher or lower expression of a few dozen genes, these differences may not be readily discovered through supervised pharmacogenomic analysis of small data sets. The modest gene expression differences between responders and nonresponders can be masked by the larger scale molecular differences due to any phenotypic imbalance between the response groups and the technical noise of microarray experiments. For example, when 24,000 measurements are performed (e.g., Affymetrix U133A gene chip) and the overall concordance is 97.98% in a technical replicate (i.e., same RNA profiled twice), 1.31% of all measurements can have twofold or greater variation. This means that 314 genes decreased or increased twofold from one experiment to another due to technical noise alone (26). If the sample size is large, this technical noise can be separated from the true signal, but when sample sizes are small it is more difficult to separate noise from the true signal.

PHARMACOGENOMIC MARKER DEVELOPMENT PROCESS

Ideally, the development process of molecular markers should resemble the development stages of therapeutic agents (Fig. 1) (27). In this context, a phase I marker discovery study would demonstrate the technical feasibility of developing a predictor and assess the reproducibility of marker results. The second phase

Development stages of pharmacogenomic predictors

<table>
<thead>
<tr>
<th>Phase I</th>
<th>Phase II</th>
<th>Phase III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery Marker optimization</td>
<td>Independent validation of accuracy</td>
<td>Prove clinical utility</td>
</tr>
<tr>
<td>LAB 99-402 Neoadjuvant biopsy study (N&gt;600)</td>
<td>MDACC 2003-0321 Randomized study (FAC vs T/FAC, n=270)</td>
<td>MDACC 2006-0543 Molecular Triaging Program</td>
</tr>
<tr>
<td>- Feasibility</td>
<td>- USO 02-103 / MDACC Phase II TX/FEX (n=100/190)</td>
<td>- MDACC 2006-0089 Personalized treatment selection for MBC</td>
</tr>
<tr>
<td>- Reproducibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Cross platform performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Signature development for T/FAC, FAC, TX/FAC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

WF Symmans et al. Cancer 2003
M Ayers et al. J Clin Onc 2004
J Stec et al. J Mol Diagn 2005
R Rouzier et al. Clin Cancer Res 2005
Y Ji et al. Applied Bioinformatics 2006
MAQC. Nature Biotechnology 2006
Y Gong et al. Lancet Onc 2007
WF Symmans et al. SABCS, 2006

Figure 1   Schema of the developmental stages of molecular markers.
includes phase II validation studies of statistically adequate power to estimate more precisely the predictive performance of the test in independent validation cases. The next level of marker evaluation would include the demonstration of clinical value by showing that a particular clinical endpoint improves when the new test is used compared with the current standard, which could be no testing or another type of testing. This may require a prospective phase III marker validation study. Currently, no pharmacogenomic response predictors have completed all three levels of clinical evaluation. Most published studies represent phase I marker discovery and phase II marker validation studies.

FEASIBILITY OF MICROARRAY-BASED DIAGNOSTIC TESTS

Technical Reproducibility

The reproducibility of microarray measurements has been examined systematically by the Microarray Quality Control Project (MAQC), which included the US Food and Drug Administration and 51 academic and industry collaborators. The goal of this project was to assess the within- and between-laboratory technical reproducibility of microarray measurements and compare the results across different microarray platforms. Four different RNA samples were profiled in five replicates on seven different array platforms, three different laboratories representing each platform. The same RNA samples were also profiled by three different reverse-transcription polymerase chain reaction (RT-PCR) methods. The most important finding of this collaborative effort was that the microarray measurements were highly reproducible within and across the different commercially available microarray platforms (28,29). The median coefficient of variation (CV) for within-laboratory replicates ranged from 5% to 15% for the various platforms and it was 10% to 20% for between-laboratory replicates. Most importantly, the concordance between the microarray based mRNA measurements and RT-PCR results were also high. The correlation coefficients for gene expression ratios measured by arrays and by RT-PCR in two different samples ranged from 0.79 to 0.92 for several hundreds of genes.

Between-platform variation in gene expression measurements is due to sequence variations in the probe sets that target the same gene at different locations. Factors that determine signal intensities generated by distinct probes for the same gene include GC content of the sequence, sequence length, cross-match opportunities, and the location of the probe sequence in relation to the 3’end of the target gene. The concordance between signal intensity for probes that target the same gene on different platforms is directly related to sequence homology (30,31). Probes with complete sequence match yield highly concordant results across platforms.

Few studies examined the within- and cross-platform reproducibility of multigene prediction scores, as opposed to the reproducibility of single gene measurements. Two reports indicated high reproducibility of pharmacogenomic
Prediction results in replicate experiments from the same RNA, when the same experimental procedures were used (26,32). Furthermore, when random noise was added to corrupt the gene expression data, a remarkable degree of robustness was observed in the face of noise.

However, caution must be exercised when trying to reproduce prediction results across different platforms. There are multiple sources of variation that will decrease reproducibility including differences in probe set and different normalization methods. Essentially, all published results that attempted cross-platform testing of gene signatures reported diminished, but not completely lost, classification accuracy on data generated by other than the original platform (33).

**Preanalytical Sources of Variation**

Another potentially important source of variation in microarray data is preanalytical variations, including differences in tissue sampling and RNA degradation. There are several different techniques to obtain tissue from a patient for pharmacogenomic studies, including excisional biopsy (with or without microdissection), fine needle aspiration (FNA), and core needle biopsy (CNB). Both needle biopsy techniques yield sufficient RNA for gene expression profiling with DNA microarrays (1–3 µg) (34–36). However, different sampling methods may result in different cellular composition and distinct gene expression profiles (37). FNA specimens contain mostly neoplastic cells with few infiltrating leukocytes, whereas the tumor cell content of CNB can vary from 10% to 90%. It is important to consider the different amount of stromal cell that might be present in a specimen, because the presence or absence of stromal genes might influence prognostic or predictive gene signatures. Also, in studies comparing pre- and postchemotherapy samples, an important source of variation may be the different cellular composition of the specimens.

RNA is an easily degradable molecule and therefore was considered suboptimal target for marker measurements compared with DNA or proteins. The development of single step universal RNA stabilizing and preserving reagents (RNA later, Ambion) has completely changed the field of RNA-based diagnostics (38). Today, RNA can be stabilized within seconds after a fresh tissue biopsy has been obtained and stored at room temperature for several days (if needed) due to readily available commercial RNA stabilizing solutions. Under these circumstances, RNA degradation from properly collected fresh tissues is now rare. Interestingly, it is now more cumbersome to preserve proteins for future analysis than it is to preserve RNA. A single step, universal protein preserving solution is yet to be discovered. However, it is important to realize that surgical resection methods (i.e., length of warm ischemia) and the length of time between removing a piece of tissue from the body and placing it into RNA preservative have a major effect on RNA quality (39).
STATISTICAL DESIGN OF PHARMACOGENOMIC DISCOVERY AND VALIDATION STUDIES

Clinical biomarker discovery and validation is governed by similar statistical rules as therapeutic clinical trials (40). Estimation of discovery sample size is essential to develop a reliable pharmacogenomic predictor, and appropriately powered independent validation studies are needed to convincingly demonstrate its predictive values. There are several statistical methods that could be used to estimate the number of patients needed to develop a clinical outcome predictor (41). One simple approach is to power the study to discover individual genes that show a predefined level of differential expression (often expressed in statistical terms as standard effect size) between outcome groups. However, the number and extent of differentially expressed genes are often unknown in advance, and the multiple comparisons (i.e., thousands of different measurements on each case) represent a major bias in this approach. Another approach takes into account accumulating preliminary results to estimate the sample size needed for discovery (42). With multivariate prediction algorithms, it is expected that the predictive accuracy increases as the predictor is developed from and trained on larger and larger data sets until it reaches a plateau (i.e., best possible predictor within the constraints of the technology). The general principle is to fit a learning curve into existing data by incrementally increasing the training set size and extrapolate results to larger sample sizes. This is an intuitively appealing approach, but it requires a substantial amount of preliminary results from many dozens of cases. It also implies that discovery sample size will vary depending on the difficulty of separating the two response groups on the basis of gene expression results. When we applied this method to human breast cancer gene expression data to estimate the number of cases needed for discovery of a chemotherapy response predictor, the results indicated that 80 to 100 cases would yield a predictor as good as extending the sample size to 200 (32).

To assess the accuracy of the predictor, it is necessary to test it on independent patient cohorts. Often, there are not enough cases available for validation, and investigators report cross-validation results. In cross-validation, some samples are left out from the marker development process and later on used to test the accuracy of the predictor, and this process is repeated many times. This leave-one-out cross-validation method gives a reliable estimate of the overall accuracy of the marker development process, but it is not the validation of a particular predictor. The genes that contribute to the predictors differ from iteration to iteration. The final single best predictor that is developed from the entire training data must be tested on completely separate cases to assess its predictive value. The goal of an independent validation study is to (i) define the true sensitivity, specificity, and the positive and negative predictive values (PPV and NPV) with narrow confidence intervals (CI) and (ii) to prove clinical utility of the test. Since the estimated performance characteristics of the test are known from the cross-validation results, the design of validation studies is more

Liedtke and Pusztai
Gene Expression Profiling to Personalize Chemotherapy Selection

straightforward. Sample size for a single arm validation studies can be calculated on the basis of the known historical response rate in unselected patients and the estimated PPV of the genomic predictor. For example, if we assume that the response rate in unselected patients is 30% and the true PPV of the test is 60% (two-fold greater chance of response in “test-positive cases”), the sample size could be calculated as follows: The lower bound of a two-sided 95% CI should not be less than 0.3 (since the expected maximum pCR rate is 30% in unselected cases). This implies that the standard error for the PPV is less than 0.1 and therefore the study must include at least 24 “test-positive” patients. The proportion of patients with response is expected to be around 30%, and an accurate prediction algorithm should generate a proportion that test positive also about 0.20 to 0.40, this equates to a total sample size of 80 to 120 patients for validation. If the predictor performs better and the true PPV is higher than 60%, fewer patients may be sufficient. A randomized trial design could simultaneously address response rates in selected and unselected patients and could also assess treatment specificity of the predictor.

BREAST CANCER PHARMACOGENOMIC RESPONSE PREDICTORS

Supervised Pharmacogenomic Predictors Developed from Human Data

Results from several phase I genomic response marker discovery studies were reported that compared gene expression profiles of cases with response and no-response, and used the genes that were differentially expressed to construct multigene predictors. Some of these markers were also assessed in small independent validation cohorts. Table 1 summarizes the currently available literature in this field. Chang et al. presented the first study that suggested the feasibility of developing multigene predictors of response to neoadjuvant chemotherapy. They developed a 92-gene predictor for clinical response to docetaxel using gene expression data from 20 patients with locally advanced breast cancer (36). Leave-one-out partial cross-validation indicated a predictive accuracy of 88% (95% CI, 68–97%), sensitivity of 85% (95% CI, 55–98%), and specificity of 91% (95% CI, 59–100%). Larger scale independent validation of this marker set has not yet been reported. Another small study examined 24 breast cancer patients treated with neoadjuvant paclitaxel, 5-fluoruracil, doxorubicin, and cyclophosphamide (T/FAC) chemotherapy and developed a 74-gene predictor for pCR (43). This multigene predictor model, tested on an independent cohort of 12 patients, showed a predictive accuracy of 78% (95% CI, 52–94%), PPV of 100% (95% CI, 29–100%), NPV of 73% (95% CI, 45–92%), sensitivity of 43% (95% CI, 10–82%), and specificity of 100% (95% CI, 72–100%). A larger follow-up study included 82 cases for marker discovery and examined 780 distinct pharmacogenomic predictors in cross-validation (32). Many of the distinct predictors demonstrated statistically equal performance in complete cross-validation. This (text continues on page 91)
<table>
<thead>
<tr>
<th>Study</th>
<th>Patient characteristics and treatment</th>
<th>Method/hypothesis</th>
<th>Gene expression platform</th>
<th>Results</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang et al., Lancet 2003 (36)</td>
<td>24 (=dev.) and 6 (=val.) pts. with locally advanced breast cancer → 4 cycles 3-weekly docetaxel</td>
<td>Development, LOOCV and independent validation of a gene expression profile predictive for clinical response (≤25% RD)</td>
<td>HgU95-Av2 (12,625 genes)</td>
<td>92 differentially expressed genes ($p &lt; 0.001$) from different functional classes; Combination into prediction model LOOCV: 88% accuracy, 92% PPV, 83% NPV, 85% sensitivity, 91% specificity</td>
<td>First study developing a multigene predictor for response to chemotherapy in breast cancer patients</td>
</tr>
<tr>
<td>Ayers et al., JCO 2004 (43)</td>
<td>24 (=dev.) and 12 (=val.) pts. with primary breast cancer → 3 cycles 3-weekly or 12 cycles weekly paclitaxel followed by 4 cycles of FAC (T/FAC)</td>
<td>Development and independent validation of gene expression predictor for pCR</td>
<td>cDNA arrays (30,721 genes)</td>
<td>Development of 74-gene-multigene predictor Independent validation: 78% accuracy, 100% PPV, 73% NPV, 43% sensitivity, 100% specificity</td>
<td>No single gene sufficiently associated with pCR to serve as single valid marker Proof of principle for building multigene predictors for chemotherapy response</td>
</tr>
<tr>
<td>Study</td>
<td>Number of Patients</td>
<td>Treatment</td>
<td>Method/Hypothesis</td>
<td>Platform</td>
<td>Results</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>-----------</td>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Iwao-Koizumi, JCO 2005 (44)</td>
<td>44 (dev.) and 26 (val.) pts. with stage II/IV breast cancer</td>
<td>→ 4 cycles of docetaxel</td>
<td>Development and independent validation of gene expression predictor for clinical response (WHO criteria)</td>
<td>ATAC-PCR (2,453 selected genes)</td>
<td>Development of 85-gene signature</td>
</tr>
<tr>
<td>Hannemann et al., JCO 2005 (45)</td>
<td>48 pts. with LABC</td>
<td>→ cycles of doxorubicin/docetaxel (AD, n = 24) or doxorubicin/cyclophosphamide (AC, n = 24)</td>
<td>Identification of gene expression predictor for pCR or nPCR for response to AD and/or AC</td>
<td>Proprietary array (18,432 genes)</td>
<td>Failure to identify differentially expressed genes or define multigene predictor</td>
</tr>
<tr>
<td>Gianni et al., JCO 2005 (46)</td>
<td>89 (dev.) and 82 (val.) pts. with LABC</td>
<td>→ 3 cycles of 3-weekly doxorubicin/paclitaxel followed by 12 cycles weekly paclitaxel</td>
<td>Identification of gene expression markers that predict response</td>
<td>RT-PCR (384 genes)</td>
<td>86 genes correlating with pCR (p &lt; 0.05) forming three clusters (ER, proliferation and immune-related gene cluster)</td>
</tr>
<tr>
<td>Patient characteristics and treatment</td>
<td>Method/hypothesis</td>
<td>Gene expression platform</td>
<td>Results</td>
<td>Comment</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>------------------</td>
<td>--------------------------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Folgueira et al., Clin Cancer Res 2005 (47)</td>
<td>Development and independent validation of gene expression predictor for pCR</td>
<td>692 cDNA microarray</td>
<td>25 differentially expressed transcripts between responders and nonresponders, 3 gene classifier could not be validated; successful development and validation (LOOCV and independent dataset) of distinct three-gene classifier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dressman et al., Clin Cancer Res 2006 (48)</td>
<td>Definition of gene signature for (a) inflammatory breast cancer (IBC), (b) tumor hypoxia, and (c) clinical response and validation in two independent datasets</td>
<td>Affymetrix HG-U133 Plus 2.0</td>
<td>Identification and validation (LOOCV) of (a) 22-gene signature for IBC (b) 18-gene signature for tumor hypoxia Failure to identify gene signature predicting clinical response</td>
<td>Development and successful validation (2 independent datasets) of a signature predictive for lymph node persistence</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Patients</td>
<td>Treatment Method</td>
<td>Hypothesis</td>
<td>Platform</td>
<td>Results</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>------------------</td>
<td>------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Thuerigen et al., JCO 2006 (49)</td>
<td>52 pts. with breast cancer (48 pts. with breast cancer)</td>
<td>5 cycles biweekly gemcitabine/epirubicin followed by 4 cycles biweekly docetaxel (GEsDoc) → 6 cycles of 3-weekly gemcitabine/epirubicin/docetaxel (GEDoc)</td>
<td>Development and independent validation of gene expression predictor for pCR</td>
<td>Oligonucleotide array (21329 genes)</td>
<td>512-gene-signature predictive of pCR with independent validation: 88% accuracy, 64% PPV, 95% NPV, 78% sensitivity, 90% specificity</td>
</tr>
<tr>
<td>Cleator et al., Breast Cancer Res Treat 2006 (51)</td>
<td>40 pts. with primary breast cancer</td>
<td>6 cycles doxorubicin/cyclophosphamide (AC)</td>
<td>Development and LOOCV of predictor for clinical response</td>
<td>Affymetrix HG-U133A</td>
<td>253 differentially expressed genes; 75 and 178 genes over-expressed in resistant and sensitive tumors, respectively</td>
</tr>
</tbody>
</table>

(Continued)
Table 1  Overview of Published Studies on Pharmacogenomic Predictors of Response to Neoadjuvant Chemotherapy in Breast Cancer (Continued)

<table>
<thead>
<tr>
<th>Patient characteristics and treatment</th>
<th>Method/hypothesis</th>
<th>Gene expression platform</th>
<th>Results</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hess et al., JCO 2006 (32)</td>
<td>Development and independent validation of gene expression predictor for pCR</td>
<td>Affymetrix HG-U133A</td>
<td>Identification of 30-gene predictor</td>
<td>Follow-up study of (43)</td>
</tr>
<tr>
<td>82 (=dev.) and 51 (=val.) pts. with primary breast cancer → 3 cycles 3-weekly or 12 cycles weekly paclitaxel followed by 4 cycles of FAC (T/FAC)</td>
<td>Validation fivefold cross-validation and permutation testing</td>
<td></td>
<td>Predictor preserved its predictive accuracy when applied to replicate samples</td>
<td></td>
</tr>
<tr>
<td>Mina et al., Breast Cancer Res Treat 2006 (52)</td>
<td>Identification of genes correlating with pCR Correlation of pCR with RS or ER-related genes Identification of genes correlating with inflammatory phenotype</td>
<td>RT-PCR assay (274 genes)</td>
<td>22 of 274 candidate genes correlated with pCR ($p &lt; 0.05$) forming three large clusters (angiogenesis, proliferation, and invasión-related genes) Lack of correlation between pCR and RS/ER-related genes</td>
<td>Largest study to date</td>
</tr>
<tr>
<td>45 pts. with stage II/III breast cancer (including 6 pts. with inflammatory BC) → 3 cycles biweekly doxorubicin and 6 cycles weekly docetaxel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ABC transporter, ATP-binding cassette transporter; dev., development/training set; IBC, inflammatory breast cancer; LABC, locally advanced breast cancer; LOOCV, leave-one-out cross-validation; npCR, near pCR; NPV, negative predictive value; pCR, pathological complete response; PPV, positive predictive value; pts., patients; RD, residual disease; RS, recurrence score (Oncotype DX®) (53); RT-PCR, reverse transcription polymerase chain reaction; val., validation set.
observation is confirmed by similar reports in the genomic literature, which indicates that several different but equally good predictors can be developed from the same data. This is due to the coordinated expression of many thousands of genes. Any combination from a particular group of coexpressed genes can carry similar information. However, a nominally best 30-probe set genomic predictor of pCR (DLDA-30) was selected for independent validation on 51 cases, and its performance was compared with a multivariate clinical variable–based predictor including age, estrogen receptor status, and grade. The genomic predictor showed a sensitivity of 92% that was significantly higher than that of the clinical predictor (61%). The NPV (96% vs. 86%) and area under the receiving operator curve (0.877 vs. 0.811) were also slightly better. Several other smaller studies demonstrated similar proof of concept that this supervised approach to marker development is feasible (Table 1).

**Genomic Predictors from Cell Line Models**

An interesting alternative approach is to use cancer models to develop chemotherapy response predictors and test the performance of these in human data. The gene expression profiles of breast cancer cell lines are different from those of human breast cancers, but nevertheless, they recapitulate many biological and genomic properties of primary breast cancers (54). Several groups reported drug-specific transcriptional changes in response to chemotherapy exposure in vitro that might be exploited as predictors of therapy response in patients (55). However, the predictive value of these in vitro generated signatures in patients has not yet been studied extensively. Nevertheless, this is an important research direction that if working can have important implications for individualizing treatment regimens for patients. In fact one could envisage custom-assembled chemotherapy regimens based on sensitivity profiles for individual drugs. It is rapidly becoming feasible to test such concepts in silico. Several groups have made their complete gene expression data from neoadjuvant biopsies publicly available. Cell line-based predictive signatures can be tested on these already existing and clinically annotated human data sets (56).

**Gene Expression Changes in Serial Biopsies**

Similar to cell lines, transcriptional response to chemotherapy was also observed in serial biopsies before and during chemotherapy. The rationale is that these changes may be more predictive of response than baseline gene expression profiles. However, the few small studies that examined this approach reported very heterogeneous transcriptional response to chemotherapy (57). In essence, different genes change in different individuals. Sotiriou et al. examined gene expression changes after one complete cycle of neoadjuvant chemotherapy and demonstrated that tissue samples, obtained from patients with good response, exhibited much greater changes in gene expression than those from poor
responders (34). These observations were corroborated by results obtained by Hannemann et al., who compared gene expression profiling on pretreatment biopsies \((n = 46)\) and posttreatment tumors \((n = 17)\) obtained at time of surgery. They also demonstrated that neoadjuvant chemotherapy causes major changes in gene expression in those cases who responded to treatment compared with those who did not (58,59). Several technical and theoretical concerns limit our enthusiasm for this approach to predictor development. It is difficult to obtain serial biopsies at regular intervals from patients who undergo chemotherapy. There is substantial variation in the actual timing of the biopsy compared with the planned time of obtaining the follow-up biopsy. For example, biopsies planned at 24 hours after chemotherapy are in fact usually collected during an 18- to 36-hour window. Since the kinetics of important gene expression changes are unknown, this introduces unpredictable variation into already highly variable data. Also, there are altogether missing time points, which further reduce the already small sample size for discovery. These factors increase the risk for false discovery from serially collected small pharmacogenomic data sets. More importantly, the practical value of a response predictor that is applied to a biopsy after chemotherapy was given is not particularly appealing.

**FUTURE POTENTIAL OF PHARMACOGENOMICS**

Pharmacogenomic predictive marker discovery is a new science. It is relatively less mature than the field of prognostic gene signatures because of the lack of preexisting tissue banks from patients who received preoperative chemotherapy. All human biopsies had to be collected in prospective clinical trials in the past few years, unlike the readily available frozen and archived tumor banks that were suitable for prognostic signature discovery and validation. However, several such data sets now exist and the field has likely entered an accelerated phase of development.

We have learned several important lessons. \(i\) It is clear that the gene expression profiles of breast cancers with extreme chemotherapy sensitivity are different from those with lesser sensitivity. Unfortunately, these gene expression differences correspond to major clinical pathological phenotypes such as ER-negative status and high grade that somewhat limits the clinical value of the current predictors. However, as larger data sets become available, it will be possible to attempt to develop predictors after stratification by known clinical pathological variables. \(ii\) It is also clear from many different reports in the genomic literature that there is no unique, single best signature for any particular classification problem. There are several distinct predictors that can perform equally well (60). \(iii\) The availability of large, public, clinically annotated, comprehensive gene expression databases started to revolutionize the biomarker discovery process. It is now possible to test in silico any mRNA-based marker for prognostic or chemotherapy or endocrine therapy predictive values, respectively, on data sets that were prospectively assembled for this purpose (61). As
more neoadjuvant data sets become available from patient cohorts treated with different treatment regimens, it will be possible to examine the regimen specificity of the various predictors. (iv) The most important long-term benefit from pharmacogenomic research may not be the current predictive signatures or the ones that will be developed in the near future, which still represent relatively crude attempts to match patients with existing drugs, but rather the large molecular data sets that this research generates. These databases could allow the identification of novel drug targets in human data and the development of drugs that target subsets of patients with particular genomic abnormalities.

REFERENCES


Gene Expression–Based Predictors of Prognosis and Response to Chemotherapy in Breast Cancer

Soonmyung Paik
Division of Pathology, NSABP Foundation, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION
With the explosion of genomic technology, numerous studies have demonstrated the potential clinical utility of gene expression profiles as a predictor of prognosis and response to therapy. Of particular interest is the demonstration that there is a general link between poor prognostic features and response to chemotherapy. In addition, linearity of RNA-based assays shows a strong promise that they can be very useful as predictive assays for targeted therapies.

Since many excellent reviews have been written on this subject, this chapter will concentrate mainly on conceptual background to aid the clinical readers understand what is going on in the field rather than going into details.

GENE EXPRESSION PROFILING PROVIDES PROGNOSTIC INFORMATION
Gene expression profiling using microarray platforms allows interrogation of expression levels of essentially all known genes in the tumor cells (1). In most studies, gross tumor rich area is homogenized to extract RNA. Therefore, the profile that is achieved is mean expression levels of each gene among different
cell population in the tumor, including host stromal cells or inflammatory infiltrates. It is also important to realize that current methods do not have sensitivity to allow absolute measurement of gene expression levels of all genes printed on the microarray. Instead only about 50% of genes will have enough signal above background level (so-called present call). Different microarray platforms use DNA probes of various lengths (ranging from 22-mer oligonucleotide in Affymetrix GeneChip, 60-mer oligonucleotides for Agilent arrays, and average 1 kilobase for cDNA arrays) representing different positions in each gene between the arrays. Sensitivity and dynamic range for measuring a specific gene could be widely different among different array platforms. Thus different studies using different platforms may result in identifying different set of genes representing same clinical phenotype (2). Obviously most studies also differ in the study population. Until recently microarray gene expression profiling required fresh or snap frozen tissue, seriously limiting the scale of studies conducted so far.

It is not entirely surprising to find gene expression profiling to provide prognostic information, since after all morphology-based prognostic factors such as tumor histological grade is essentially end result of gene expression changes in the tumor cells. Two seminal papers using microarray gene expression profiling, one by Sorlie et al. (3) and the other by van ’t Veer et al. (4), clearly demonstrated the association of gene expression profiles and prognosis of patients diagnosed with early breast cancer. However, these two studies used very different approaches. In the seminal paper published in 2000, Perou et al. used unsupervised analysis of gene expression data and identified distinct biological subtypes (normal-like, luminal A and B, basal-like, and HER2) (5). Subsequently Sorlie et al. demonstrated that these subtypes confer differing prognosis with basal-like and HER2 associated with poor clinical outcome (3). On the other hand, van ’t Veer et al. conducted a supervised analysis to develop prognostic index based on expression levels of 70 genes identified from microarray analysis (4). This was then validated in a larger cohort (6). Recently Dr. Perou’s group has demonstrated that there is a significant overlap of information between the two approaches using same dataset—meaning those classified as good prognosis using 70-gene van ’t Veer set was found to be largely populated by luminal A subtype (2).

While studies described above had enormous impact on how we view breast cancer, despite recent clearance of 70-gene Mammaprint™ assay by the U.S. Food and Drug Administration (FDA), actual clinical implementation has been impeded by the need for fresh or snap frozen tumor tissue because of the requirement for highest quality RNA as a starting material for the microarray gene expression profiling.

DEVELOPMENT AND VALIDATION OF OncotypeDx™
RECURRENT SCORE ASSAY

RNA isolated from formalin-fixed, paraffin-embedded tissue (FFPET) samples is a poor material for gene expression profiling. Masuda et al. analyzed the RNA extracted from FFPET for chemical modifications (2,7). In freshly made FFPET
that were fixed and processed in ideal conditions (fixed in 10% buffered formalin at 4°C), RNAs are fairly well preserved. Although the extracted RNAs did not show any sign of degradation compared with fresh samples, they were poor substrates for cDNA synthesis and subsequent polymerase chain reaction (PCR) amplification, so that only PCR amplification of short targets was possible. The investigators found addition of monomethylol (CH₂OH) groups to all four bases to varying degrees and some adenines dimerized through methylene bridging. In addition to the chemical modification, however, RNAs in FFPE T continue to be degraded or fragmented over time during storage for reasons unclear. Cronin et al. systematically examined the quality of RNA extracted from breast cancer FFPE T specimens taken at different times. RNA from FFPE T archived for approximately 1 year were less fragmented than RNA archived for nearly 6 or 17 years (8).

There have been reports on the use of reverse transcription polymerase chain reaction (RT-PCR) to measure expression levels of few genes using RNA isolated from FFPE T, but large-scale RT-PCR was not reported until recently. Cronin et al. from Genomic Health Inc. has invested on developing high quality controlled robotic assays for RT-PCR and published proof of concept study (8). In the latter study, while RNAs isolated from old formalin fixed tissue were heavily fragmented with average size under 200 base pairs and absolute signal was much less for older samples, normalization of signal using a set of reference genes could largely correct these problems.

Encouraged by these results, the National Surgical Adjuvant Breast and Bowel Project (NSABP) has engaged in a collaborative study with Genomic Health Inc. to develop prognostic gene expression assay for node negative (N−) estrogen receptor–positive (ER+) breast cancer treated with tamoxifen (9). While there are many other outstanding clinical questions that need to be addressed, we were compelled to first address the question of who among N− ER+ tamoxifen treated patients need chemotherapy. In addition, there were two independent tamoxifen treated cohorts (tamoxifen arms of B-14 and B-20 trials) in the NSABP tissue bank so that discovery and validation could be performed without difficulty. We wanted to make sure that by the time of reporting, the assay was validated in a completely independent cohort. At that time, technology and cost limited Genomic Health investigators to examine only 250 genes maximum in RNA extracted from 60 μm thick section from a paraffin block. Therefore, existing cancer literature and microarray data were surveyed and 250 candidate genes were selected. Before tumor blocks from B-20 tamoxifen treated cohort were examined, two other cohorts, one from mixed patient population and the other from those with more than 10 positive nodes were examined with the same assay (10). We did not believe that there will be different prognostic genes based on nodal status and therefore selected the prognostic genes that were equally prognostic across all three cohorts as a basis for prognostic model building. From the beginning, because of the continuous nature of RNA expression data, we believed strongly that a linear prediction model needed to be developed.
Correlating clinical outcome with expression levels yielded many genes from these three studies, and 16 top-performing genes were identified for final model building and validation. Relative expression levels of the 16 genes are measured in relationship to average expression levels of 5 reference genes. While a majority of genes comprising 16 genes are ER (ER, PGR, BCL2, SCUBE2) and proliferation (Ki67, STK15, Survivin, CCNB1, MYBL2) related, there are other genes (HER2, GRB7, MMP11, CTSL2, GSTM1, CD68, BACG1) (9). The unscaled recurrence score (RSu) was calculated with the use of coefficient that are defined on the basis of regression analysis of gene expression. Recurrence score (RS) was rescaled from the unscaled recurrence score (Rsu) as follows: $RS = 0$ if $Rsu < 0$; $RS = 20 \times (Rsu−6.7)$ if $0 \leq Rsu \leq 100$; and $RS = 100$ if $Rsu > 100$ (9). This resulted in $RS$ ranging from 0 to 100.

Final validation of the $RS$ was achieved by examination of its performance in a completely independent cohort from NSABP trial B-14, which was not used in the model building process (9). The validation study was conducted with rigorous predefined statistical analysis plan with prespecified outcome end points and cutoffs for $RS$. The predefined low-risk group ($RS$ below 18) had significant better prognosis compared with higher risk groups. More importantly, as shown in the published report, there is a linear relationship between $RS$ and risk of distant failure at 10 years among N– ER+ tamoxifen-treated patients (9).

To test the applicability of $RS$ in community oncology setting, Habel et al. conducted a nested case-control study in breast cancer patients diagnosed from 1985 to 1994 at 14 Kaiser hospitals (11). Eligibility included negative nodes, age less than 75 years, and no chemotherapy. Cases died of breast cancer prior to 2002. Up to three controls were matched to each case on age, race, tamoxifen treatment, facility, diagnosis year, and follow-up time. Among 4,964 potentially eligible patients, 220 cases and 570 matched controls were identified. Median age was 59 years (range, 28–74 years); 30.9% were tamoxifen treated (mostly after 1988). Median follow-up time was 4.9 years for cases (time to death) and 12.9 years for controls. $RS$ was significantly associated with breast death in ER+ patients treated with or without tamoxifen ($p = 0.002$). This study is important since it provides validation of $RS$ assay used in a community setting.

**PREDICTION OF RESPONSE TO CHEMOTHERAPY**

Among gene expression–based markers, only $RS$ has been tested whether it predicts response to chemotherapy in adjuvant setting (12). Since many of the genes comprising 21 genes in the $RS$ assay are ER or proliferation related, we hypothesized early on that $RS$ would be predictive of response to chemotherapy.

Existence of chemotherapy arm in trial B-20 allowed us to ask this question (in fact this exactly was the reason why we first decided to work with N– ER+ tamoxifen-treated cohort since we knew that we could not only discover and validate the prognostic index, but also test whether it predicts response...
to chemotherapy using banked materials readily available to us) (12). There were 651 evaluable patients (227 randomized to tamoxifen and 434 randomized to tamoxifen plus chemotherapy). Patients with tumors who had high RS greater than 30 had a large absolute benefit of chemotherapy [with an absolute increase in distant recurrence free survival (DRFS) at 10 years of 27.6 ± 8.0%, mean ± SE]. Patients with tumors who had low RS less than 18 derived minimal, if any, benefit from chemotherapy (with an absolute increase in DRFS at 10 years of −1.1 ± 2.2%, mean ± SE). The test for interaction between chemotherapy treatment and RS was statistically significant ($p = 0.038$). Of particular importance is the finding that there is a linear relationship between RS and degree of benefit from chemotherapy when RS was examined as a continuous variable in Cox Model.

While the relationship between RS and chemotherapy response has not been validated in an independent cohort as was done for prognostic aspect of the assay, supportive evidences exist to allow its use in clinical decision making. The most important data are from a cohort of patients treated with neoadjuvant chemotherapy at Milan reported by Gianni et al. (13). In this cohort, higher RS correlated with higher incidence of complete pathological response (pCR). In addition, due to the fact that many genes in RS are proliferation and ER related, we can make reasonable assumption that B-20 data are true. Since the decision to get chemotherapy has to be based on evaluation of baseline risk and assessment of the degree of potential benefit from adding chemotherapy, it is reasonable to conclude that those with lower RS may not need chemotherapy since their baseline risk is low on the basis of B-14 data, and expected benefit is low on the basis of B-20 data. One headache is the fact that because of the continuous nature of the RS, it is difficult to pinpoint from which RS score patients start to gain benefit from chemotherapy. Clinical oncologists are not used to this kind of data. Uncertainties regarding the degree of benefit from chemotherapy in patients with intermediate range of RS resulted in launching a large randomized clinical trial named Trial Assigning IndividuaLized Options for Treatment (TAILORx) by North American Breast Cancer Intergroup (14). Many questioned about the reasons for the definition of the intermediate risk group in TAILORx trial being different from the definition used in the original publication in B-14 cohort. This was due to the fact that primary clinical end point in TAILORx is disease-free survival instead of distant disease-free survival and consideration of 95% confidence intervals in the baseline risk assessment. Again one has to realize that these cutoffs are totally arbitrary.

**PREDICTION OF RESPONSE TO NEOADJUVANT CHEMOTHERAPY**

Preoperative systemic therapy trials provide unique opportunities to assess sensitivity of tumor cells to systemic therapy in vivo while treatment is ongoing. Numerous studies have demonstrated the correlation of pCR and long-term outcome in breast cancer patients, potentially eliminating the need for long-term
followup. Together with ease of obtaining fresh pretreatment core biopsies allowing high throughput assessment of molecular markers, preoperative trials may allow speedy discovery and validation of predictive markers for response to systemic therapy. So it is not surprising that many investigators attempted to utilize neoadjuvant setting to develop predictive markers of response to chemotherapy using microarray-based gene expression profiling. Regrettably the cumulative experience has so far been somewhat disappointing. No profile has been found to have enough accuracy to be directly clinically applicable, although Hess et al. demonstrated that gene expression-based predictor still provided better prediction than clinical predictors (15). But more importantly, there has been no study with large enough sample size to provide conclusive answers as to whether one could reliably predict pCR using gene expression levels.

It probably is somewhat naïve to expect gene expression profile alone to predict pCR. For example, one would expect a smaller tumor would have higher chance of getting into pCR than a bigger tumor with same gene expression phenotype. Host factors such as drug metabolism and distribution may influence the chance for pCR. Prediction of combination regimens also may be more difficult than predicting response to single regimen. However, in the latter situation, because of low pCR rate, it is very difficult to develop predictor unless sample size is very high. In this regard, perhaps what investigators have achieved so far is as far as they can get, unless these other factors are integrated into algorithm development. In addition, pCR itself is not a perfect measure of response either. As many as 80% of tumors show clinical response on taxane-based neoadjuvant chemotherapy, and therefore among those not achieving pCR, there is a continuous spectrum of residual burden in the breast. So it may make more sense to correlate gene expression with residual tumor burden than pCR, which has a somewhat arbitrary cutoff.

Despite these limitations, certain consensus has emerged that are in agreement with what we found in B-20 study—that the tumors with poor prognostic signatures have higher chance of responding to chemotherapy. In addition to above described study by Gianni et al. (13), Rouzier et al. (using gene expression array) (16) and Carey et al. (using IHC surrogates of intrinsic subtypes) (17) has examined whether Perou’s intrinsic subtypes correlate with chance of pCR. As expected from B-20 data, basal and HER2 subtypes have higher chance of pCR compared to luminal subtypes. So there is a general agreement in the field that genetic program that determines the prognosis is tightly linked with general sensitivity to chemotherapy.

**ASSESSMENT OF RESIDUAL RISK AFTER CHEMOTHERAPY**

Despite the fact that gene expression data alone have not been shown to be able to predict pCR with high accuracy, gene expression profiling can be still important in patients treated with neoadjuvant chemotherapy. In an unpublished study, using pretreatment core biopsy samples from NSABP B-27, we found that
the combination of prognostic profile with pCR may be used to define residual clinical risk after chemotherapy, since those with pCR had a very good prognosis despite having a poor prognostic gene signature. This means that patients with poor prognostic gene signature who did not achieve pCR are at high risk of recurrence after chemotherapy. Such data will allow identification of patients who will require more than chemotherapy without the need for waiting until they develop recurrences and opens door for a concept for post-neoadjuvant targeted therapy trials.

**CLINICAL UTILITY OF RS ASSAY**

**How Is RS Compared to NCCN or St. Gallen Clinical Practice Guidelines?**

In B-14 study, less than 10% of the tamoxifen treated patients are classified as low risk by either NCCN or St. Gallen clinical guidelines. These patients did have excellent prognosis with less than 10% distant failure within 10 years. In contrast, RS classified 51% of the same cohort into low-risk category with very similar 10-year failure rate. Therefore 40% more ER+ tamoxifen-treated patients would be spared from chemotherapy if RS is used in comparison to NCCN or St. Gallen clinical guidelines. In addition, unlike NCCN and St. Gallen guidelines, which assumed that treatment benefit is equal in all subgroups, NSABP B-20 data showed that the benefit from chemotherapy is nearly zero in low-risk group defined by RS.

**How Is RS Compared to Adjuvant on Line Algorithm?**

The high cost of the RS assay (estimated at $3400) can only be justified by the additional information that it provides compared with other assays. Bryant and colleagues, in collaboration with Peter Ravdin, compared RS with Adjuvant on Line (AOL) in the NSABP B-14 cohort in a presentation made at the 2005 St. Gallen Breast Cancer Symposium (personal communications with the late Dr. John Bryant). The AOL program provides individualized estimates of recurrence risk on the basis of information such as patient age, tumor size, nodal involvement, and histological grade. It is actually not that straightforward to compare RS and AOL, since AOL does not provide risk scores. Therefore, it is difficult to utilize the numerical output from AOL when comparisons are made to RS. To compare the two, Dr. Bryant rank ordered output from AOL so that similar proportion of patients belong to low- (50%), intermediate- (25%), and high-risk (25%) categories as defined by RS. Rank-ordered AOL (R-AOL) performed remarkably well as a prognostic factor (in B-14) as well as predictive factor for benefit from chemotherapy (in B-20). RS and R-AOL were both independently prognostic in B-14 in multivariate analysis, suggesting that both gene expression signature and clinical features influence patient prognosis. However, agreement between the two assays was only 48%. RS was able to tease
out patients with poor prognosis from the R-AOL low-risk category, so that the 10-year distant relapse rate was 5.6% for patients categorized as low risk by both R-AOL and RS. R-AOL low-risk but RS intermediate- and high-risk patients had a 12.9% distant relapse rate at 10 years in the tamoxifen-treated cohort ($p = 0.04$). For R-AOL intermediate- or high-risk patients, the distant relapse rate was 8.9% when the RS risk was low, whereas the distant relapse rate was 30.7% when the RS risk was intermediate or high ($p < 0.0001$). These data suggest that the RS adds to existing prognostic markers and it would be ideal to incorporate RS and AOL into a single algorithm. However, at this point we don’t have a clear way to actually implement this finding into AOL in clinical use.

**Applicability in Node-Positive Patients**

One question about RS is whether it will work for N$^+$ ER$^+$ patients. There is no reason why it should not work when it actually works even in 10$^+$ node patients. However, risk estimates for RS will be different between N$^-$ and N$^+$ patients although overall direction of the data would be the same, so one cannot simply apply the current RS algorithm to N$^+$ patients without a study.

**Categorical Vs. Continuous Variable**

One of the biggest problems in the field currently in my view is the use of RS as a categorical variable (low-, intermediate-, and high-risk categories), which was arbitrarily defined for validation study to ease the understanding of the data, because clinicians are not used to seeing a continuous data (9). It is very important to recognize the continuous nature of the RS. For example, a patient with score 17 (categorized as low risk) has nearly identical risk as a patient with score 19 (categorized as intermediate risk). On the other hand, a patient with RS of 3 would have much different prognosis than the one with RS of 17, although they are both categorized as low risk in the published report.

**FUTURE PERSPECTIVE**

The problem of current gene expression profiling methods is the need for fresh or snap frozen tissue and/or high cost. There are alternative methods that may be able to overcome these barriers. For example, Quantiplex assay (Panomics) based on combination of branched DNA amplification and fluorescence-labeled beads may allow multiplexing as many as 30 genes at a time in a 96 well plate with reasonable amount of input RNA extracted from FFPET. cDNA-mediated annealing, selection, extension, and ligation (DASL) assays from Illumina is another promising example (18). These assays cost fraction of the cost for RT-PCR or microarray-based methods and can utilize FFPET as starting materials. Replacement of standard markers such as ER and HER2 by these assays or QRT-PCR is desirable since they will provide more linear prediction of response.
Availability of gene expression profiling methods for FFPET opens the possibility that archived materials from adjuvant chemotherapy trials can be analyzed, and therefore current approach of using gene expression analysis in neoadjuvant setting as a mean to identify predictive markers will become less important. The latter approach always will require validation in adjuvant setting since the FDA will not approve marker based on pCR as a surrogate—therefore it makes more sense to perform discovery and validation of predictive markers purely in the adjuvant setting.

REFERENCES

Utilization of Genomic Signatures for Personalized Treatment of Breast Cancer

P. Kelly Marcom, Carey K. Anders, Geoffrey S. Ginsburg, and Anil Potti

Duke Institute for Genome Sciences & Policy, Department of Medicine, Duke University Medical Center, Durham, North Carolina, U.S.A.

Joseph R. Nevins

Duke Institute for Genome Sciences & Policy, Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina, U.S.A.

INTRODUCTION

The challenge of personalized cancer treatment lies in matching the most effective therapy to the characteristics of an individual patient’s tumor. The selection of therapy for both early- and late-stage breast cancer, as is the case for most cancer therapy, is still largely empiric and guided by large, population-based, randomized clinical trials. Although widely adopted, this approach is inadequate for the selection of individualized chemotherapy regimens. Estimates of benefit are extrapolations from the effects observed in these large trials and are not necessarily applicable to individual patients.

The use of genomic data, in particular gene expression profiles derived from DNA microarray analysis, has the promise of providing the basis for truly individualized therapy. This approach has the potential to transform the treatment of
cancer from a population-based approach to an individualized one directed toward unique patient characteristics. This includes the use of gene expression data to identify subtypes of cancer not previously recognized by traditional methods of analysis. Profiling genomic patterns to identify new subclasses of tumors has previously been described in distinguishing differences between acute myeloid leukemia and acute lymphoblastic leukemia (1,2). Similar studies have employed gene expression data to identify distinct subtypes of breast cancer with prognostic significance (3).

A landmark study in breast cancer published in 2002 described the use of a DNA microarray-based, 70-gene expression profile as a prognostic factor in breast cancer (4). This genomic marker has demonstrated predictive value stratifying stage I and II breast cancer patients into two broad subgroups, relatively high-risk versus low-risk groups as defined by a 10-year risk of breast cancer recurrence. Various other studies have described the development of expression profiles that have prognostic value in breast cancer (5–9).

The Netherlands 70-gene predictor has provided the basis for an expanded European study that aims to measure its performance in a more diverse set of patients (10). The parallel, prospective clinical trial MINDACT (Microarray In Node-negative Disease may Avoid ChemoTherapy) aims to measure the effectiveness of this genomic predictor in guiding adjuvant chemotherapy when compared to predictions based solely on the traditional clinicopathologic features. Likewise, a large prospective trial in the United States, TAILORx (Trial Assigning Individualized Options for Treatment), is currently evaluating the Oncotype DX test (Genomic Health, Inc., Redwood City, California, U.S.), a polymerase chain reaction (PCR)-based assay of 21 genes correlated with 10-year breast cancer recurrence risk (11), to guide the administration of hormonal therapy with or without chemotherapy among patients with estrogen receptor (ER)-positive, early-stage breast cancer.

Of equal, if not greater, importance in achieving the goal of personalized treatment of breast cancer is the ability to predict response to specific chemotherapeutics, particularly the standard-of-care cytotoxic regimens incorporated into routine clinical practice. Various studies have described approaches to developing genomic predictors of response to chemotherapeutic agents promising a more effective strategy of assigning chemotherapeutic regimens to patients, as guided by unique characteristics of their breast tumors (12–14).

The overall aim of this chapter is to highlight both opportunities presented by recent advances in the development of genomic-based approaches within the field of breast cancer and corresponding strategies for implementation of these techniques into clinical practice. Rather than relying on historic paradigms of chemotherapeutic selection as guided by broad, population-based studies, the goal of this unique methodology is to improve patient outcome, and ultimately patient survival, through the optimal selection of patient-specific chemotherapeutic agents as guided by genomic-based strategies.
CYTOTOXIC THERAPIES FOR EARLY-STAGE BREAST CANCER

Currently, the standard regimens for the adjuvant treatment of early-stage breast cancer combine anthracyclines, taxanes, and cyclophosphamide in limited permutations of combination, dose, and administration schedule. The addition of a taxane, either sequentially or in combination, to a core anthracycline/cyclophosphamide regimen in the setting of node-positive breast cancer has been shown to improve disease-free survival (DFS) by 4% to 7% in four separate randomized trials (15–18). However, it is unclear whether all breast cancer subtypes are equally affected by this therapeutic approach. Recent work examining a series of breast cancer adjuvant trials conducted over the past two decades further confirms the importance of defining breast cancer subtypes and their relationship to treatment benefit. This work demonstrated a lack of benefit for this broad treatment approach across a subset of ER-positive patients (19).

Optimizing chemotherapy outcomes through the manipulation of dosing schedules remains controversial. In CALGB 9741 (Cancer and Leukemia Group B Intergroup 9741), decreasing the dosing interval from every three weeks to every two weeks [i.e., dose-dense (DD) scheduling] was shown to improve DFS and overall survival (OS) at three-year median follow-up [risk ratio (RR) = 0.74, \( p = 0.010 \)] (20). However, an update at five-year median follow-up showed a less impressive overall improvement in DFS (RR = 0.80, \( p = 0.012 \)) and perhaps no improvement in ER-positive disease (21). Moreover, a randomized Phase III trial from the Gruppo Oncologico Nord Ovest–Mammella Inter Gruppo (GONO–MIG) examining DD scheduling of a non-taxane-containing regimen did not show a benefit in DFS or OS (22). In an accompanying editorial to this publication, the ambiguity regarding the benefit of DD therapy was acknowledged and emphasis was placed on the need to define biologically relevant breast cancer subtypes so that individualized treatment strategies could be developed (23).

A ROLE FOR PREOPERATIVE SYSTEMIC CHEMOTHERAPY TO EVALUATE GENOMIC STRATEGIES

Preoperative, systemic chemotherapy (PST), also known as neoadjuvant therapy, is an effective means for assessing the sensitivity of an individual breast tumor to chemotherapy. Multiple trials of preoperative systemic chemotherapy have been performed. Initially, this approach was primarily used for locally advanced, inoperable, or inflammatory breast cancer. However, subsequent trials have evaluated PST in the setting of smaller, operable breast tumors both to increase breast conservation rates and to assess sensitivity to chemotherapy.

Over the past two decades, PST for early-stage breast cancer has been shown to be a safe and effective method for assessing the chemosensitivity of primary breast cancers. Sequencing systemic therapy first, with the consequent
three- to six-month delay in surgery, has not been shown to adversely affect outcomes (24). Assessment of pathologic response to PST, both in the breast and regional lymph nodes, provides the most powerful prognostic information available. Patients who obtain a complete pathologic response in the breast and lymph nodes have a greater than 95% survival at seven years (24,25). Although PST provides a powerful tool for in vivo assessment of chemosensitivity, exploiting the information gained in the aforementioned trials to improve management of individual patients has been more difficult.

Studies employing a PST approach have yielded conflicting results illustrating the heterogeneity of primary breast tumors and responsiveness to different chemotherapeutic agents. This observation begs the question on whether or not breast tumors commonly possess intrinsic ‘pan-chemotherapy’ responsiveness or resistance. For example, patients responsive to preoperative treatment with CVAP (Cytoxan, vincristine, Adriamycin, and prednisone) were either continued on the same regimen or crossed over to docetaxel (26). Responding patients who received docetaxel had a significantly higher pathologic complete response (pCR) rate than those who continued on the CVAP regimen, supporting differential breast tumor sensitivity to anthracyclines and taxanes. Additionally, the pCR rate in the NSABP (National Surgical Adjuvant Breast and Bowel Project) B-27 clinical trial was doubled by the addition of preoperative docetaxel to preoperative doxorubicin/cyclophosphamide (AC) chemotherapy (26.1% vs. 12.9%, \( p < 0.0001 \)) supporting this conclusion. However, the addition of preoperative docetaxel did not yield a statistically significant difference in overall DFS when added to preoperative AC alone [hazard ratio (HR) = 0.9, \( p = 0.22 \)]. The question remains open on whether patients only partially responsive to AC, who achieved a pCR following administration of docetaxel, might have achieved a pCR with docetaxel alone, thus avoiding the risk of cardiotoxicity inherent to anthracyclines. Unfortunately, this study is somewhat flawed by the concurrent administration of tamoxifen with chemotherapy in patients with known ER-positive breast tumors (25). An additional study by the GEPARDUO (German Preoperative Adriamycin Docetaxel) study compared DD doxorubicin/docetaxel (AT) with AC followed by docetaxel, illustrating improved response among patients receiving cyclophosphamide. This study, however, was more widely interpreted as a comparison of sequential versus DD therapy (27).

The largest trials investigating response to PST with AC come from the NSABP B-18 and the aforementioned B-27 clinical trials, where the pCR rates, allowing for residual in situ disease, were consistently approximately 13% (25,28). Data for single-agent paclitaxel are more limited; however, a study conducted at the M.D. Anderson Cancer Center reported pCR rates of 6% when excluding in situ diseases (29). A comprehensive review of these and other PST trials supports that the pCR rate in the breast with most doublet chemotherapies is approximately 10% to 15% (24). The National Comprehensive Cancer
Network (NCCN) guidelines support that standard adjuvant regimens can also be used as preoperative therapy (http://www.nccn.org).

Although response to PST is a powerful tool for assessing tumor chemosensitivity, it is a post hoc evaluation and requires the exposure of individual patients to toxic and potentially ineffective therapies. Consequently, it has been difficult to exploit the knowledge gained about both tumor sensitivity and resistance to the patient’s benefit. In addition, to our knowledge, no studies have examined the influence of PST on the postoperative management of patients who have minimal or no response to PST.

The relative importance of including either an anthracycline or taxane or both as part of the adjuvant treatment regimen of early stage breast cancer is an area of active research. These two classes of agents, in combination with cyclophosphamide, provide the foundation for the treatment of early-stage breast cancer, and several clinical trials provide insight into the current use of these agents. A recently published randomized phase III trial compared AC with docetaxel/cyclophosphamide (TC) (30). While this trial demonstrated improvement in DFS favoring TC (86% vs. 80%, \( p = 0.015 \)), the confidence interval (CI) approached 1.0 (CI = 0.50–0.94) and there was no difference in OS despite median follow-up of 5.5 years. Interestingly, a previous trial comparing AT with AC showed no difference in DFS at four years (31). Although both of these trials enrolled a lower-risk population of patients, reconciling differences in outcome across these two studies is difficult. An ongoing trial in patients with fewer than four positive axillary nodes is comparing AC with single-agent paclitaxel as adjuvant therapy (CALGB 40101), in part on the basis of responses observed in the aforementioned PST trials. In summary, although many active chemotherapeutic agents suitable for the treatment of early-stage breast cancer exist, data available in 2007 do not provide the necessary guidance to select the most effective chemotherapeutic agent for each unique patient. This observation speaks to the importance of developing novel methods employing genomic analyses of tumor biology to individualize and optimize patient care.

GENE EXPRESSION PROFILES TO PREDICT THERAPEUTIC RESPONSE

Given the current status of neoadjuvant and adjuvant chemotherapy for breast cancer, an opportunity exists to use genomic information to improve the selection of patients who would derive the greatest benefit from treatment with cytotoxic chemotherapy. That is, prognostic tools that can identify patients clearly in need of further treatment can help to refine the decisions based now on clinical information. Additionally, the development of predictive signatures specific for response to individual cytotoxic agents would provide an invaluable clinical tool to enable the better use of available standard-of-care agents. Historically, predictive biomarkers in the setting of breast cancer have played an important role in selecting patients for targeted therapies (i.e., trastuzumab and endocrine therapies
including tamoxifen and aromatase inhibitors). However, validated biomarkers for commonly used cytotoxic agents are not yet clinically available. The importance of cytotoxic agents, even with the advent of the targeted molecular therapy era, cannot be discounted. To date, many targeted therapies demonstrate greatest efficacy when combined with chemotherapeutic agents. To address this relevant clinical need, a variety of studies have been performed with the aim of developing predictors of sensitivity for a spectrum of cytotoxic chemotherapies.

Gene expression profiling of breast cancers in the context of preoperative chemotherapy has shown promise in providing predictive signatures of response to specific agents (12–14). The availability of clinical datasets provides an opportunity to move ahead with the prospective assessment of gene expression profiling as a means for guiding PST. This methodology can also simultaneously provide insight into breast cancer biology, as the investigation of profiles will shed light on the deregulation of molecular pathways.

Several studies have correlated breast cancer gene expression profiles with response to preoperative chemotherapy. Chang et al. performed gene expression profiling on locally advanced breast cancers treated with docetaxel to identify genes predictive of response to this therapeutic modality (32). Similar work has been done with preoperative doxorubicin therapy and sequential taxane/anthracycline therapy (33,34). While some studies have suggested that single markers such as NF-κB may predict resistance to anthracycline therapy, most have found that a multigene classifier is needed. In comparing a signature for doxorubicin/cyclophosphamide sensitivity with a signature for docetaxel sensitivity, Cleator et al. concluded that unique signatures for sensitivity to different chemotherapy regimens are likely to exist (33).

More recently, studies of PST have examined the pCR rates for clinically and molecularly defined breast cancer subtypes. These retrospective analyses of PST trials demonstrate marked differences in pCR rates depending on breast cancer subtype (35,36). A recent publication by the M.D. Anderson Breast Group has retrospectively defined gene expression profiles that predict response to a sequential regimen of paclitaxel (Taxol) followed by 5-fluorouracil, doxorubicin (Adriamycin), and cyclophosphamide (T-FAC) (13).

As an alternate strategy, drug sensitivity data from the NCI-60 panel of cancer cell lines, coupled with baseline Affymetrix gene expression data unique to these cells, has been utilized to generate a series of gene expression signatures with the potential to predict sensitivity to various chemotherapeutic agents (37) (Fig. 1). To test the capacity of the in vitro docetaxel sensitivity predictor to accurately identify the patients who responded to docetaxel, we evaluated data from a previously published neoadjuvant study that linked gene expression data with clinical response to docetaxel (32). The in vitro–generated profile correctly predicted docetaxel response in 22 of 24 patient samples, achieving an overall accuracy of 91.6%. Findings were further validated in an ovarian cancer dataset, where accuracy for predicting the response to salvage docetaxel therapy exceeded 85% (37).
Given that the in vitro developed gene expression profiles predicted clinical response to docetaxel, the approach has been extended to test the ability of additional signatures to predict response to commonly used salvage therapies in an independent dataset of ovarian cancer patients previously treated with Adriamycin. As shown in Figure 1B, each of these predictors was capable of accurately predicting the response to the drugs in patient samples, achieving an accuracy in excess of 81% overall. Importantly, in these examples, the overall clinical response rate varied from 25% to 45%. Using the expression signatures to predict patients likely to respond would essentially increase the “effective” response rate (the positive predictive value for chemosensitivity for any given drug) to greater than 85%, by selecting those patients likely to respond, based on the drug sensitivity predictor. These data suggest that the genomic-based predictors of chemotherapeutic response provide a unique opportunity to guide the selection of a cytotoxic agent most effective for an individual patient—the essence of personalized medicine.

Importantly, it is also evident from further analyses that while there are overlaps in the predicted sensitivities to the chemotherapeutic agents among ovarian patients, there are also distinct groups of patients that are predicted to be sensitive to various single-agent salvage therapies (Fig. 2). Many therapeutic regimens make use of combinations of chemotherapeutic cytotoxic agents, raising a question on the extent that signatures of individual therapeutic response will also predict response to a combination of agents. To address this question,
data from a breast neoadjuvant trial were utilized to evaluate the ability of single-
agent signatures for T-FAC to predict patient response measured by pCR (34). Utilization of individual chemotherapy sensitivity signatures to assess patient response illustrated a significant distinction between the responders \((n = 13)\) and nonresponders \((n = 38)\), with the exception of 5-flourouracil. Importantly, the combined probability of sensitivity to all four agents in the T-FAC neoadjuvant regimen was calculated using the probability theorem. It was clear from this analysis that the prediction of response based on a combined probability of sensitivity, built from individual chemosensitivity predictions, yielded a statistically significant \((p < 0.0001, \text{Mann–Whitney U test})\) distinction between the responders and nonresponders.

**A STRATEGY FOR USING CHEMOTHERAPY RESPONSE PREDICTORS IN PST**

Although promising, studies to date have included post hoc gene expression profiling of tumor samples. The ability to effectively select the most appropriate preoperative systemic cytotoxic agents for the treatment of early-stage breast cancer must ultimately be tested and validated in prospective clinical trials. This rationale approach will not only provide an opportunity to confirm biologically relevant subtypes, but also establish their predictive capacities regarding response to specific, conventional cytotoxic agents.

---

**Figure 2 Patterns of predicted sensitivity to commonly used chemotherapies.** A collection of breast cancer samples represented by microarray data was used to predict the likely sensitivity to the indicated agents based on the signatures at the top of the figure. Predictions are plotted as a heatmap representing probability of sensitivity. Adapted from Potti et al., Nature Medicine 12:1294–1300, 2006.
The development of genomic signatures predictive of response to commonly employed chemotherapeutics offers an opportunity to significantly expand the identification of combinations of cytotoxic agents best suited for an individual patient. With this in mind, several prospective clinical studies have been proposed to test the capability of genomic predictors to most effectively match cytotoxic agents with individual patients. This concept is best evaluated in the context of neoadjuvant trials where hypotheses regarding the effectiveness of genomic predictions can be accurately assessed through pathologic measurements of tumor response.

We note that the availability of chemotherapy response predictors provides an opportunity to employ genomic predictors in present-day practice, where patients are commonly treated with one or more therapeutic regimens of equal efficacy (Fig. 3). The ability to select the most effective therapy, among a panel of standard-of-care agents, is a rational strategy because standard treatment procedures would not be greatly altered. The goal of these trials would be to more effectively utilize currently available standard agents. For example, a breast neoadjuvant study has been designed and is currently being implemented by employing this approach. Patients with early-stage breast cancer will be treated with one of two current standard-of-care therapies for which genomic predictors are available. This study design would compare the response rates for the randomly or doctor-selected A or B regimens versus assignment to therapy A or B via genomic predictor of response. The success of this study could then set the stage for larger studies utilizing genomic predictors of chemotherapy sensitivity to more accurately predict response to cytotoxic therapies. The next phase of clinical studies would then focus on novel therapeutics (i.e., targeted therapies) for the individual patients unlikely to respond to either regimen.

Figure 3  Schematic diagram of prospective clinical trial that would evaluate the capacity of predictors of chemotherapy response to improve response rates.
OTHER GENOMIC OPPORTUNITIES TO GUIDE BREAST CANCER THERAPEUTICS

The successful development and validation of predictors of response to standard chemotherapeutics will foster the more efficient use of these agents through the identification of patients with the greatest likelihood of response. Concurrently, these strategies will also identify a subgroup of patients resistant to all available standard-of-care chemotherapeutics, necessitating the implementation of targeted strategies. This is a critical point as it is of little benefit to patients to identify resistance if a plan for alternative therapeutic strategies does not exist. Currently, most patients resistant to standard chemotherapeutic agents are traditionally treated with second- or third-line chemotherapeutic agents or are enrolled on a clinical trial. The more effective alternative is to identify patients likely to be resistant prior to initiation of therapy so that opportunities for alternative treatments might be employed in the first-line setting.

A viable option is to exploit the wealth of pathway-specific therapeutics that have been developed over the past two decades. To this goal, gene expression patterns reflecting deregulation of oncogenic pathways have been developed, as depicted in Figure 4 (38). These oncogenic pathway signatures have been shown to accurately predict the status of pathway deregulation in a series of both tumors derived from murine models and human tumors. These results demonstrate an ability to predict pathway profiles unique to individual human tumors, thus characterizing the genetic events associated with tumor development. Importantly, the assay of gene expression profiles provides a measure of the consequence of the oncogenic process, irrespective of how the pathway might have been altered. Thus, even if the known oncogene is not mutated, but rather another component of the pathway is altered, the gene expression profile will detect this alteration.

The analysis of oncogenic pathways through gene expression profiling offers an opportunity to identify new therapeutic options for patients by providing a potential basis for guiding the use of pathway-specific drugs. To demonstrate the value of pathway prediction to guide drug selection, pathway deregulation was mapped among a series of breast cancer cell lines to screen potential therapeutic drugs. In parallel with pathway status mapping, the cell lines were treated with a variety of drugs known to target specific activities within given oncogenic pathways. A variety of available reagents known to target oncogenic pathways for which signatures have been developed [i.e., a farnesyl transferase inhibitor (FTS) and an Src inhibitor (SU6656)] were employed. In each case, a clear relationship was demonstrated between prediction of pathway deregulation and sensitivity to the respective targeted therapeutic. These preliminary, yet promising, results have been further validated in a larger set of 50 cancer cell lines, including lung, ovarian, and melanoma tumor types.

Taken together, these data demonstrate a potential approach to identify therapeutic options for chemotherapy-resistant patients, as well as identify novel
combinations for chemotherapy-sensitive patients. This unique approach represents a potential strategy to optimize effective treatments for cancer patients and warrants future, prospective, validations trials in both the adjuvant and metastatic settings.

Pathway signatures have the potential to identify patients most sensitive to standard-of-care cytotoxic drugs, as well as those resistant to standard regimens. This approach creates an opportunity to combine chemotherapy sensitivity profiles with concurrent oncogenic pathway deregulation within an individual patient’s tumor. As depicted in Figure 5, predictions of sensitivity to

Figure 4 Patterns of predicted cell signaling pathway activation. A collection of breast cancer samples represented by microarray data was used to predict the likely state of activation of the indicated cell signaling pathways based on the signatures at the top of the figure developed to represent pathway activation. Predictions are plotted as a heatmap representing probability of pathway activation. A Kaplan-Meier survival analysis is shown at the bottom reflecting the overall survival for patients in each cluster defined by pathway patterns. Source: Adapted from Bild et al., Nature 439:353–357, 2006.
various cytotoxic agents can be combined with predictions of oncogenic pathway deregulation with the goal of defining the most optimal regimen. For example, a patient whose tumor is resistant to docetaxel may concurrently illustrate a high probability of Src pathway deregulation. Src inhibitors are currently being tested in clinical trials in a variety of solid tumors, including breast tumors. The selection of a therapy targeted against the Src pathway may prove more efficacious for this individual patient compared with the empiric selection of a cytotoxic agent. In addition, this strategy could be employed to identify novel combinations of therapies—either multiple cytotoxic agents or cytotoxic agents combined with targeted agents—as guided by a more rational and informed approach.

**IMPACT AND FUTURE**

The practice of oncology continually faces the challenge of matching the optimal therapeutic regimen with the most appropriate patient, balancing relative benefit with risk, to achieve the most favorable outcome. This challenge is often daunting with marginal success rates in many advanced disease contexts, likely reflecting the enormous complexity of the disease process coupled with an inability to properly guide the use of available therapeutics.

The importance of selecting patients who will respond to a given therapeutic agent is perhaps best illustrated by the example of trastuzumab. In the absence of selection as guided by Her2 protein overexpression or gene amplification, the overall response rate among breast cancer patients is on the order of
10%. In contrast, for patients selected on the basis of Her2 overexpression or amplification, the overall response rate when combined with chemotherapy approached 50% to 60% (39–41). The use of trastuzumab in the treatment of both early-stage and advanced breast cancer is guided by Her2 protein overexpression of gene amplification. Gene expression signatures predicting response to various cytotoxic chemotherapeutic agents provide an equally important opportunity to optimize response to various cytotoxic agents.

Importantly, clinically-available, validated predictors of chemotherapy response could provide an invaluable tool in present day practice. Commonly, patients are treated with one or more therapeutic regimens that, on a population basis, have individually demonstrated equal efficacy. Validated chemotherapy predictive of chemotherapy sensitivity provides an opportunity to select the most effective therapy among a panel of standard-of-care agents individualized to a particular patient’s tumor biology. It is essential that this approach be incorporated into modern day clinical trials. Ultimately, this novel concept has the potential to augment future treatment strategies incorporating likelihood of sensitivity to common cytotoxic agents, as well as various targeted therapeutic agents. Incorporating this approach into the routine management of patients with various malignancies has the potential to more rationally assign combination therapies, both cytotoxic and targeted agents, as rationally guided by each individual’s unique tumor biology.

SUMMARY

Treatment within the field of oncology, in particular that of breast cancer, is shifting away from a broad, population-based approach toward individualized care as driven by unique tumor biology. Validated signatures of chemotherapy sensitivity, coupled with an understanding of oncogenic pathway signaling within a given breast tumor, provides an opportunity to improve the utilization of standard-of-care therapies. This approach also provides an opportunity to take the next step toward personalized cancer care—incorporation of targeted agents—as guided by advances in genomic technology.

REFERENCES


Genomic Signatures for Personalized Treatment of Breast Cancer


INTRODUCTION

For accurate management of early breast cancer, several important decisions need to be made by patients and clinicians regarding the type of surgery, the need for adjuvant radiotherapy, and particularly the type of adjuvant systemic therapy (chemotherapy and/or endocrine therapy). The decision about adjuvant systemic therapy is of paramount importance since breast cancer is mainly a systemic disease, with a majority of deaths occurring as a consequence of distant metastases. This decision depends on the likelihood of recurrence; patient-related factors such as biological age, menopausal status, performance status, and co-morbidities; and tumor-related factors such as size, grade, lymph node status, and hormonal and human epidermal growth factor receptor 2 (HER2) status. Most of these variables are combined into prognostic models or guidelines such as the St. Gallen consensus or the National Comprehensive Cancer Network (NCCN) guidelines, in order to assist patients and clinicians in the decision-making process (1–3). Although these guidelines provide valuable information about prognosis in certain breast cancer populations, the prediction of disease
outcome for the individual patient is far from accurate. In reality, patients with similar clinicopathological characteristics can have strikingly different outcomes. Because of these limitations, and because of the recognition of the incurable nature of metastatic breast cancer, clinicians consider prescribing adjuvant chemotherapy in the majority of early-stage breast cancer patients to reduce the risk of relapse. However, a large proportion of these patients is probably cured with locoregional treatment alone or in combination with adjuvant endocrine therapy, thus unnecessarily exposing many patients to chemotherapy side effects. The identification of robust and reliable prognostic markers that accurately select patients not requiring adjuvant chemotherapy is essential to decrease the problem of overtreatment. Additionally, a more accurate patient selection will also contribute to the decrease in undertreatment, a less common but dangerous problem.

The introduction of high-throughput techniques, such as microarray-based gene expression profiling and sequencing of the human genome, has led to a novel approach in breast cancer research. Measuring the expression level of thousands of genes in one experiment using microarray technology enables us to better understand the biology and heterogeneity of breast cancer. The first study to examine gene expression patterns in breast cancer showed the existence of at least four molecular subtypes of breast cancer—luminal A, luminal B, basal, and HER2 positive—distinguished by extensive differences in gene expression (4). Several subsequent studies confirmed these findings (5–8). In addition, microarray techniques have been used to identify profiles that correlate with disease outcome (prognostic profile) or with response to treatment (predictive profile) (9–16).

One of the first prognostic profiles was the 70-gene profile (MammaPrint™) identified by van ’t Veer et al. (13). This dichotomous microarray classifier can accurately distinguish breast cancer patients who have a high likelihood of remaining free of distant metastases (good profile) from those who are at high risk of developing distant metastases (poor profile) (Fig. 1).

**DEVELOPMENT AND FIRST RETROSPECTIVE VALIDATION OF THE 70-GENE PROFILE (MAMMAPRINT™)**

The 70-gene prognostic profile was developed in a training set of 78 lymph node–negative, invasive breast tumors less than 5 cm in diameter. Patients were younger than 55 years at diagnosis, had no prior malignancies, and were treated in the Netherlands Cancer Institute (NKI) with breast conserving therapy or mastectomy. Five out of 78 patients received adjuvant systemic treatment comprising chemotherapy \( (n = 3) \) or endocrine therapy \( (n = 2) \). Thirty-four patients developed distant metastases within five years of diagnosis, while the remaining 44 patients remained free of distant metastases. The mean follow-up of patients who remained free of distant metastases was 8.7 years, and the mean time to metastases was 2.5 years.

For all 78 samples, the percentage of tumor cells in a hematoxylin- and eosin-stained section, before and after cutting sections for RNA isolation, was at
RNA was isolated and labeled with a fluorescent dye. Reference RNA consisted of an equal amount of RNA from all tumors. Both sample RNA and reference RNA were hybridized on an oligonucleotide microarray platform containing approximately 25,000 genes (produced by Agilent Technologies, Palo Alto, California, U.S.) (17). After hybridization, microarray slides were washed and scanned using a confocal laser scanner (Agilent). Fluorescence intensities were quantified, corrected for background noise, and normalized (13).

Using a statistical method called “supervised classification,” 231 genes that were differentially expressed in tumors that metastasized compared with tumors that did not were selected. These 231 genes were then ranked on the basis of their correlation coefficient with disease outcome. Using a leave-one-out cross-validation procedure, the top 70 of these genes appeared to be the optimal set of genes to accurately discriminate between tumors with a good disease outcome (free from distant metastases for at least five years) and tumors with a poor disease outcome (distant metastases within five years of diagnosis).

All 78 tumors were ranked according to their correlation with the average expression of the 70 genes of the patients who did not develop distant metastases. The sensitivity was optimized by setting a threshold, which resulted in the...
misclassification of 3 out of 34 patients with a poor disease outcome being classified as good prognosis and therefore would erroneously have had chemotherapy withheld (9% misclassification). Subsequently, the 70-gene profile was validated in 19 breast cancer tumors—7 tumors from patients with a good clinical outcome (no distant metastases within five years of diagnosis) and 12 tumors from patients who did develop distant metastases within five years of diagnosis. Using the 70-gene profile, only 2 out of 19 tumors were classified incorrectly, thereby confirming the initial results (13,18).

The same group of investigators performed a first retrospective validation study using a consecutive series of 295 breast tumors from NKI (144 lymph node–positive and 151 lymph node–negative tumors). All patients were younger than 53 years at diagnosis and were treated with locoregional therapy alone (56%) or in combination with adjuvant systemic therapy comprising chemotherapy alone (31%), hormonal therapy alone (7%), or a combination (7%). In this consecutive series of 295 patients, 61 were also part of the previous series used to develop the 70-gene profile (14).

All 295 tumors were classified as good or poor profile by calculating the correlation coefficient with the established average expression level of the 70 genes in the previously studied good-outcome group. For the 234 samples that were not included in the previous study, tumors with a correlation coefficient above the threshold of 0.4 were classified as good profile; for the 61 tumors that were also part of the previous study, the threshold was adapted to 0.55 to correct for overfitting (18).

The 70-gene profile accurately classified 115 tumors (39%) as a good-prognosis group (good profile) with a 10-year overall survival of 95% (±2.6%) and 180 tumors (61%) as a poor-prognosis group (poor profile) with a 10-year overall survival of 55% (±4.4%). Interestingly, the 70-gene profile was not only associated with disease outcome in the lymph node–negative patients [hazard ratio (HR) for distant metastases 5.5; 95% confidence interval (CI) 2.5–12.2; \( p < 0.001 \)], but also strongly associated with disease outcome in the group of 144 lymph node–positive patients (HR for distant metastases 4.5; 95% CI 2.0–10.2; \( p < 0.001 \)). Furthermore, the only significant independent factors for prediction of distant metastases were a poor-prognosis profile, a larger tumor, the presence of vascular invasion, and no chemotherapy treatment. The 70-gene profile was by far the most powerful predictor of distant metastases, with an adjusted HR of 4.6 (95% CI 2.3–9.2; \( p < 0.001 \)) (14,18).

**THE PROGNOSTIC VALUE OF THE 70-GENE PROFILE IN A CLINICAL CONTEXT**

To assess the potential added value of this prognostic tool in a clinical context, the 70-gene profile was compared with the St. Gallen consensus guidelines or the NIH criteria used at that time (3,19). In comparison with these guidelines, the 70-gene profile assigned more patients to the good-prognosis group than the St. Gallen consensus guidelines or the NIH criteria (40% vs. 15% and 7%,
respectively). Moreover, patients identified by the 70-gene profile as having good prognosis were more likely to remain free of distant metastases than patients identified as having good prognosis according to the St. Gallen consensus guidelines or the NIH criteria. On the other hand, patients who were assigned to the poor-prognosis group by the 70-gene profile had a higher risk of developing distant metastases as compared with the poor-prognosis group classified by the St. Gallen consensus guidelines or the NIH criteria.

The 70-gene profile was built to have a minimum of misclassified patients with a poor disease outcome; in other words, the low-risk group should be maximized without causing considerable undertreatment. Consequently, in this retrospective validation series, the 70-gene profile had a high negative predictive value of 96% for the 234 new patients and a positive predictive value of only 38%. That is to say, of all patients with a good prognosis profile, 4% will still develop distant metastases and would erroneously have had chemotherapy withheld, whereas all patients classified as having a poor prognosis, 38% will indeed develop distant metastases. Consequently, treatment decisions based on the 70-gene profile would still lead to overtreatment; but since the total proportion of patients identified as having a poor profile is much smaller than the proportion of high-risk patients according to the St. Gallen consensus guidelines or the NIH criteria, overtreatment would be reduced by 25% to 30%. Furthermore, the overall prediction of disease outcome and consequent treatment decision according to the 70-gene profile seems to be more accurate.

INDEPENDENT RETROSPECTIVE VALIDATION BY THE TRANSBIG CONSORTIUM

To further substantiate the prognostic value of the 70-gene profile, the European Union’s sixth Framework Project TRANSBIG (http://www.breastinternational-group.org/TransBIG.aspx) conducted an independent retrospective validation study (20). Frozen tumor material and clinical data were collected from 302 patients from five different cancer centers in the United Kingdom, Sweden, and France. Patients were younger than 61 years at diagnosis, had a lymph node-negative T1 or T2 tumor, and had not received adjuvant systemic therapy. The patients were treated before 1999, and the median follow-up in this series was 13.6 years. The frozen tumor samples were sent to Agendia, the microarray facility in Amsterdam, where RNA was isolated and hybridized on a custom-designed microarray, measuring the gene expression level of the 70 genes in triplicate. Tumors were classified as poor profile if the correlation coefficient of the average expression was under 0.4. The microarray facility had no access to the clinical data, and the researchers collecting the clinical data were blinded for the 70-gene profile results. Only an independent statistical office had access to both the clinical and the microarray data simultaneously. Eighty percent of the samples were centrally reviewed for estrogen receptor (ER) status and tumor grade, and the centers were independently audited for quality of data control.
This independent study confirmed that the 70-gene profile can accurately select patients who are at high risk of developing distant metastases (poor prognosis) from patients who are at low risk of developing distant metastases and hence can be safely spared chemotherapy (good profile), with HRs of 2.79 (95% CI 1.60–4.87) and 2.32 (95% CI 1.35–4.0) for distant metastases–free survival and overall survival, respectively. Moreover, patients with a good profile had a 10-year overall survival of 89% compared with 69% in patients with a poor profile. The 70-gene profile outperformed several prognostic methods such as the Nottingham Prognostic Index and the St. Gallen consensus guidelines. The performance of the 70-gene profile was also compared with Adjuvant! Online. The Adjuvant! Software—available at www.adjuvantonline.com—calculated a 10-year overall survival probability based on clinicopathological criteria (the patient’s age, tumor size, tumor grade, ER status, and comorbidity) (21,22). Patients were considered as having low clinical risk when the 10-year breast cancer–specific survival probability calculated by Adjuvant! was more than 88% if their tumors were ER positive and more than 92% if they were ER negative. The rationale for these two cutoffs was the assumption that patients with ER-positive tumors would now receive hormonal therapy, with an average absolute benefit of 4%. In the discordant cases, the 70-gene profile provided much stronger prognostic information, with almost identical survival rates for the low and high clinical risk groups within each 70-gene profile risk, suggesting that the 70-gene profile predicts disease outcome independently of clinical and pathological characteristics (Fig. 2) (20).

Figure 2 Kaplan–Meier curve by clinical and 70-gene expression signature risk groups. Overall survival. Source: From Ref. 20. By permission of Oxford University Press.
Another important finding of this validation study is the time dependency of the 70-gene profile. The median follow-up in this series was twice as long as the follow-up of the initial series. To determine the influence of the time of distant metastases on the prognostic value of the profile, HRs were calculated with arbitrary censoring of all observations at different time points. The adjusted HR for distant metastases at 5 years was 4.7 in comparison with an adjusted HR of 3.5 at 10 years, suggesting a better prediction of early (within five years of diagnosis) metastases by the 70-gene profile (18,23,24).

OTHER PROGNOSTIC PROFILES IN BREAST CANCER

Several other prognostic gene expression signatures were developed, such as the 76-gene profile, the wound signature, and the genomic grade index (15,25,26). The 76-gene prognostic profile was developed on a different microarray platform (Affymetrix, Santa Clara, California, U.S.) and using a slightly different methodology; the genes associated with relapse of the disease were selected separately for ER-positive and ER-negative patients. The genes selected in each group were then combined in a 76-gene profile (15). A retrospective validation study confirmed the prognostic value of this profile in early-stage breast cancer patients (27). Recently, the 76-gene profile has also been validated independently by the TRANSBIG consortium in the same patient series, using the same methodology as the validation of the 70-gene profile (24). Although only three genes were common to both profiles, the biological pathways represented are the same, and the prognostic value of both the 70-gene profile and the 76-gene profile was validated in this patient series (20,24). Since the profiles perform equally, and the 70-gene profile has proven to be reproducible and robust and is available as a Food and Drug Administration (FDA)-cleared diagnostic test (28), the TRANSBIG consortium decided to use the 70-gene profile in a large prospective clinical trial, the Microarray In Node-negative Disease may Avoid ChemoTherapy (MINDACT) trial.

IMPLEMENTATION OF THE 70-GENE PROFILE: THE RASTER TRIAL AND THE LOGISTICS PILOT STUDY

In addition to scientific evidence provided by validation studies, implementation of these microarray profiles in both prospective clinical trials and in daily practice requires logistical feasibility. The requirement of fresh tissue as the source of high-quality RNA has several consequences for tumor sample handling. Traditional fixation of fresh tissue in formaldehyde, for example, causes degradation of RNA, thereby making tumor samples unsuitable for microarray experiments. In addition, to avoid interlaboratory variability, fresh tumor samples have to be shipped to one microarray facility. Consequently, for the implementation of microarray profiles, local procedures have to be adapted, and close collaboration between pathologists, surgeons, and medical oncologists is of utmost importance.
Recently, two multicenter feasibility studies were conducted, the RASTER study and a logistics pilot study (18,29). The first study was coordinated by NKI with financial support from the Dutch Health Care Insurance Board. The aim of the study was to assess the feasibility of collecting good-quality tissue from several community hospitals in the Netherlands to perform the 70-gene profile test. For all patients included in this study, tumor samples from excised specimens were obtained and placed in a commercially available preservation fluid (RNAlater®, Qiagen GmBH, Hilden, Germany) at room temperature. Subsequently, the samples were sent by conventional mail to NKI, where they were snap-frozen and stored at –80°C. The 70-gene profile was performed at the microarray facility in Amsterdam. Preliminary results showed that it is feasible to adapt local procedures and to collect good-quality tissue for microarray testing in a multicentric setting (18,29,30).

In the RASTER study, tissue was preserved temporarily in RNAlater and sent by conventional mail at room temperature. Preservation of tumor tissue in RNAlater may influence several processes in the tissue, such as protein levels. Since one of the additional aims of the MINDACT trial is the establishment of a biological materials bank for future research, including proteomics, storage of tissue in RNAlater is not suitable for the MINDACT trial and fresh, frozen tissue will be mandatory (31). Therefore, a second feasibility study was conducted to test the collection, storage, and shipment of fresh, frozen tissue in six European hospitals. Women younger than 71 years, with early-stage breast cancer, were included. After surgery, the pathologist obtained representative tumor samples using a 6-mm biopsy puncher. Tumor samples were snap-frozen in liquid nitrogen and stored at –80°C within one hour of surgery. The samples were shipped to Amsterdam on dry ice by a contracted courier specializing in the transportation of frozen tissue. At Agendia, RNA was isolated when the samples were representative of the tumor (i.e. tumor cells ≥ 50%). After measurement of the quality and quantity of the RNA, the samples were hybridized on the 70-gene profile platform. Results showed that in general it is feasible to collect good-quality fresh, frozen tissue for microarray tests in a multicentric and multinational setting. Additionally, in the course of this feasibility study, procedures were adapted and optimized for use in the MINDACT trial (18).

PROSPECTIVE VALIDATION OF THE 70-GENE PROFILE IN THE MINDACT TRIAL

Despite the completion of all necessary technical validation procedures and the FDA clearance as the in vitro diagnostic multivariate index assay (IVDMIA), including approval of clinical utility, the clinical validation of the 70-gene profile can still be further optimized in currently diagnosed, adjuvantly treated populations to achieve level-1 evidence. The TRANSBIG consortium has decided to further validate the 70-gene prognostic alongside the classical clinicopathological factors (31). This trial, the MINDACT trial [EORTC (European...
Organization for Research and Treatment of Cancer) 10041 BIG (Breast International Group) 3-04], will enroll 6000 women between 18 and 71 years of age, with primary invasive breast cancer at diagnosis. Tumors should be unilateral, stage T1, T2, or operable T3; carcinomas in situ are eligible provided invasive cancer is present. Patients should be treated with breast conserving treatment or mastectomy and should have a negative sentinel node or axillary clearance. Patients with primary malignancies or previous chemotherapy or radiotherapy will not be included. The main goal of the trial is to prove that the 70-gene profile will safely assign chemotherapy to a smaller proportion of node-negative breast cancer patients. Therefore, all patients included in MINDACT will have a risk assessment based on the 70-gene profile and according to the currently used clinicopathological criteria. This last risk assessment is provided by an updated version of Adjuvant! Online. If, according to both risk assessments, the patients’ risk of developing distant metastases is low (an estimated 13% of all patients), no chemotherapy will be given. If both methods classify the patients’ risk of developing distant metastases as high (an estimated 55% of the patients), chemotherapy will be proposed. If the risk assessments are discordant (an estimated 32% of all patients), patients will be randomized to the clinicopathological criteria or the 70-gene profile to be used for adjuvant chemotherapy decision making (Fig. 3). The group of most interest to answer the primary research question comprises patients with a low-risk 70-gene profile (good profile) and high-risk

Figure 3  MINDACT study design. Source: From Ref. 18. With permission. Abbreviations: R-T, treatment decision randomization; CT, chemotherapy.
clinicopathological criteria who will be randomized to use the 70-gene profile for the treatment decision and thus receive no chemotherapy. In this group a null hypothesis of a five-year distant metastases–free survival of 92% will be tested (31).

Additional study objectives are related to the type of chemotherapy and endocrine therapy. Patients who are to receive chemotherapy may be randomized between an anthracycline-based regimen and a regimen of docetaxel and capcitabine in order to compare the efficacy and safety of both regimens. The latter regimen may have increased efficacy with less long-term side effects such as cardiotoxicity and leukemia. Post-menopausal (spontaneous or induced) patients with hormone receptor–positive tumors will be randomized for endocrine therapy with two years of tamoxifen, followed by five years of letrozole or seven years of letrozole upfront.

In addition to these three research questions, whole genome arrays will be performed for all 6000 patients, and frozen and paraffin-embedded tumor samples together with blood samples will be collected from all patients. These microarray data and the biological materials will be stored in an independent biobank and will provide a valuable source of information for future research that may result in the identification of new prognostic or predictive biomarkers and new drug targets.

FUTURE PROSPECTS

The MINDACT trial is the first randomized clinical trial that will prospectively evaluate the prognostic value of a microarray-based prognostic tool, the 70-gene profile or Mammaprint, in an adjuvantly treated population. The implementation of molecular diagnostics based on the true biology of a patient’s tumor will greatly expedite the introduction of personalized medicine.

The details of the rationale, design, and logistics of the MINDACT trial are described in references 18, 23, and 31, and the authors strongly recommend these articles to the reader.

ACKNOWLEDGMENTS

The studies mentioned in this chapter were supported by the European Commission Framework Programme VI, the Center of Biomedical Genetics, the Dutch Health Care Insurance Board, the Dutch National Genomic Initiative: Cancer Genomics Program, the Breast Cancer Research Foundation, the EBCC-Breast Cancer Working Group–association sans but lucratif (asbl), the Fondation Belge Contre Le Cancer, the S. G. Komen for the Cure Foundation, Jacqueline Seroussi Memorial Foundation for Cancer Research, and the European Organisation for Research and Treatment of Cancer (EORTC)–Breast Cancer Group. S. Mook was partially supported by the traineeship program of the TRANSBIG network. The authors thank the numerous individuals who have contributed
Microarray Gene Expression Profiling

133

to the studies mentioned in this review, especially those from the TRANSBIG consortium and the EORTC, and all the patients who have and still are participating in these studies.

CONFLICTS OF INTEREST

Dr. L. J. van ’t Veer is a named inventor on a patent application for Mammaprint and reports holding equity in Agendia BV.

REFERENCES

Predictive Markers for Targeted Breast Cancer Treatment

Hans Christian B. Pedersen and John M. S. Bartlett

Endocrine Cancer Group, Cancer Research Centre,
Western General Hospital, Edinburgh, United Kingdom

INTRODUCTION

Currently selection of breast cancer treatment is guided predominantly by patient prognosis using classical pathological assessment of tumors, with higher risk patient groups being offered more aggressive therapy. The choice of treatment regimen is guided by a very small number of predictive biomarkers (Table 1) (1). However, it is now clearly recognized that not only may the risks of treatment outweigh the benefits in some patient groups but that not all patients are equal with respect to their response to and benefit gained from exposure to novel and established therapies. There is increasing interest in the identification of biological markers to improve the targeting of established therapies to those cancers likely to respond. In addition, development of new molecular targeted drugs has led to an urgent need for research into predictive markers that can distinguish patients who benefit from treatment. New proteomic technologies are leading to markers being discovered at an increasing rate, with protein expression, cellular localization, and posttranslational modification being recognized as those of potential importance. Many new immunohistochemical (IHC) markers are

---

Hans Christian B. Pedersen is employed on a Marie Curie Transfer of Knowledge project between Dako A/S, Glostrup, Denmark, and University of Edinburgh.
Currently being tested for predictive potential for various therapies in breast cancer. Through large-scale gene expression studies using microarrays, it has been realized that tumors have molecular signatures that may determine response to treatment and patient prognosis. Many candidate-predictive and prognostic signatures in breast cancer are being explored, and some are already in use to aid patient management (e.g., OncoType Dx\textsuperscript{1}, Genomic Health, Redwood City, California, U.S.). The challenge for the future is to identify both the biological and diagnostic determinates of treatment response in order to allow matching of appropriate treatments to specific patients.

Future breast cancer treatment is likely to be guided by predictive markers along with genomic signatures, treating patients with mixtures of molecular targeted therapeutics. This chapter will review current IHC and fluorescent in situ hybridization (FISH)-based predictive markers in breast and the challenges associated with the development and implementation of predictive markers in future.

**CURRENT PREDICTIVE MARKERS IN BREAST CANCER: LESSONS LEARNT**

A predictive marker is defined as a factor that indicates sensitivity or resistance to a specific treatment (2). In contrast, a prognostic marker predicts a tumor’s future behavior independent of systemic adjuvant treatment. Some markers can be both predictive and prognostic, as exemplified by the estrogen receptor \( \alpha \) (ER\( \alpha \)) expression, which both predicts response to endocrine treatment and correlates with prognosis.

**Hormonal Therapy and Hormone Receptors**

ER\( \alpha \) and progesterone receptor (PR) status were the first predictive markers used in the clinic to select patients for breast cancer treatment. Today, IHC ER\( \alpha \) or PR staining is done routinely on patient tumor samples to predict response to endocrine therapy. More correctly, the assay is used to exclude ER\( \alpha \)- or PR-negative patients from ER\( \alpha \)- or PR-targeted therapy, as they will have a low probability of responding to treatment (3). Historically however, ER\( \alpha \) was first discovered in the late 1960s on its ability to bind radiolabeled ligand (4,5).
Within 10 years of its initial discovery, assays had been developed that allowed ER to be used as a marker for hormonal responsiveness and clinical aggressiveness of the tumor (6). Many years passed before it was realized that only ERα-positive patients could benefit from treatment and for these finding to be applied as clinical practice (3). It is now almost universally accepted that ERα testing has a central role in breast cancer management, but it is worth noting that no prospective trial has yet tested the predictive value of ERα or PR as predictive biomarkers. Later evidence has emerged supporting testing for PR as well as ERα (7). PR is a transcriptional target of ligand-bound ERα, and it has been proved that including the PR score in the ERα score provides stronger evidence for true ERα positivity and therefore makes patients more likely to respond to therapy (8–10). However, controversy remains over the value of supplementary markers to ERα, and in several countries, PR testing is not routinely performed.

The drugs administered to ERα- or PR-positive patients have been through many generations through the last three decades, from pure ERα-antagonist (fulvestrant) to ERα-partial antagonist tamoxifen and drugs inhibiting the enzyme aromatase that synthesise oestrogen (11). Current clinical practice in breast cancer management utilizes a combination adjuvant tamoxifen followed up by aromatase inhibitors (AIs) or AIs alone. However, ER or PR do not select between these therapies, and despite evidence that PR is predictive of poor response on tamoxifen, there is no evidence, to date, that AIs circumvent this resistance.

Despite the long history of the ERα or PR status as a predictive and prognostic marker, no standardized protocols for fixation and IHC staining are yet widely applied. Conflicting evidence exists for when patients are considered ER negative and therefore will not respond to treatment, but above 10% positive-staining tumor cells are generally considered ER positive (12). Perhaps because ERα has been available for testing as a predictive marker for such a long time, before emphasis was put on standardizing IHC testing, this has lead to the development of a multitude of different protocols using a variety of antibodies. ER IHC results have been shown to be highly influenced by the efficiency of the antigen-retrieval step (13). External quality assurance initiatives such as the United Kingdom National External Quality Assessment Scheme for Immunocytochemistry (UK NEQAS-ICC) exist now to improve quality across clinical laboratories through external assessment of analytical and interpretative performance; such schemes are however not uniformly successful.

In addition to standardizing ERα IHC for scoring tumors as positive or negative, it has been suggested that quantification of ERα levels may be necessary in some clinical settings because tumors expressing low or intermediate ERα levels have been shown to benefit from chemotherapy in addition to endocrine therapy (14). Current ERα IHC protocols are semiquantitative, so alternatives such as radiolabeled ligand-binding assays or quantitative polymerase chain reaction (qPCR) have been suggested. Gene expression of ER and PR correlate well with protein expression, and it has been suggested to be an alternative to IHC as it is sensitive and quantitative over a large dynamic range.
range (15). It has been suggested that more quantitative methods of measuring ER and PR may provide additional information to guide choice of therapy. We will discuss this in detail later (12).

HER2/neu and Trastuzumab (Herceptin®)

HER2/neu or human epidermal growth factor receptor 2, a product of the proto-oncogene HER2 first discovered in 1985, belongs to the HER family of tyrosine kinase receptors (16–18). Numerous studies have documented HER2/neu-positive breast cancers display more aggressive disease and shortened disease-free survival (19). Using recombinant technologies, trastuzumab (Herceptin®, Genentech, South San Francisco, California, U.S.), a monoclonal immunoglobulin G1 class, humanized murine antibody, was developed to specifically target patients with breast cancer that overexpressed the HER2/neu protein (20,21). Clinical trials demonstrated the efficiency of trastuzumab for treating metastatic breast cancers and subsequent trials in combination therapy (21–25) for HER2 positive disease. Again, HER2/neu, like ER, is more predictive of which patients will not respond to treatment.

As in the case for endocrine therapy, most predictive marker development followed drug development, leading to some confusion. In initial clinical trials, patient selection was based on an IHC assay using the mouse monoclonal antibody that had been humanized to create trastuzumab (26,27). Today most patients are selected for trastuzumab therapy on the basis of HER2/neu expression levels in tumor cells (Fig. 1). In 2001, HER2/neu testing achieved standard care of practice in the American Society of Clinical Oncology (ASCO) breast cancer clinical practice guidelines. In most laboratories, tumors are screened with FDA-approved IHC tests, HercepTest® (Dako, Glostrup, Denmark) or HER2/neu (CB11) PATHWAY® (Ventana Medical Systems, Inc., Tucson, Arizona, U.S.). Ambiguous cases are then analyzed with FDA-approved HER2/neu FISH from Vysis (Downers Grove, Illinois, U.S.), Dako, or Ventana. Alternatively, in some laboratories, FISH is the only method used for HER2/neu testing. HER2/neu testing for trastuzumab therapy has undergone considerable scrutiny in numerous studies, because evidence emerged indicated that the IHC tests generated considerable amount of false positives and negatives (28–31). This discrepancy is due in part to lack of adherence to protocols, variable fixation methods, and subjectivity related to tumor scoring.

Only 20% to 35% of HER2/neu positive tumors respond to treatment, depending on the context the drug is given in, and studies have demonstrated that this may be attributed to a combination of factors, including lack of pathway dependence, receptor being inactive despite overexpressed, or the induction of resistance through various mechanisms (32). Controversy persists with regard to the optimal method for HER2/neu testing and interpretation of results (33). It has been reported that it may be more cost effective to test all tumors with HER2/neu FISH, avoiding false positives and false negatives (28). This is a possibility, as
FISH and, more critically, chromogenic in situ hybridization (CISH) technology become more widely distributed, and an agreement can be made on the best choice of HER2/neu testing.

**DEVELOPMENT OF PREDICTIVE MARKERS**

The successful targeting of HER2/neu with trastuzumab inspired the development of a wave of new therapeutic antibodies and kinase inhibitors that were aimed at inhibiting growth factor–regulated signaling. However, repeating the success of HER2/neu testing for trastuzumab therapy has proven difficult.

Inhibitors of mammalian Target Of Rapamycin (mTOR, e.g., temsirolimus, Torisel\textsuperscript{R}, Wyeth Pharmaceuticals, Inc., Madison, New Jersey, U.S.), epidermal growth factor receptor (EGFR, e.g., cetuximab, Erbitux\textsuperscript{R}, ImClone Systems, Inc., Branchburg, New Jersey, U.S.) and vascular endothelial growth factor receptor (VEGFR, e.g., bevacizumab, Avastin\textsuperscript{R}, Genentech) have shown modest or no results in clinical trials as single agent therapies in breast cancer (34–36). This may be partly attributed to lack of predictive markers of response in combination with drugs simply not being effective as single agents. However, this failure may also
reflect poorly designed approaches toward the discovery of predictive biomarkers. Predictive markers for EGFR-targeted therapy have been in focus for the past five years, and we will therefore briefly go through the history of EGFR-targeted therapy and what has been learnt for future predictive marker development. In addition, a model for codevelopment of predictive markers in relation to traditional drug development is presented.

EGFR—A Cautionary Tale

EGFR or HER1 belongs to the same family of tyrosine kinase receptors as HER2/neu and EGFR is considered an attractive target for therapeutic intervention in many cancer forms due to the receptor being frequently mutated and overexpressed (34). EGFR overexpression is associated with reduced disease-free survival and resistance to chemotherapy (37). EGFR inhibitors (e.g., erlotinib and gefitinib) and therapeutic antibodies against the extracellular domain of the receptor (e.g., cetuximab, matuzumab, and panitumumab) are currently in clinical trials or late-stage development (38). Most clinical trials have investigated EGFR inhibitors in cancers of the lung and colon, and cetuximab has been approved for treatment of colorectal cancer (34). Clinical trials have also been conducted in breast cancer, with limited success (39–47). In several of these trials, patients have been selected for trial on the basis of EGFR overexpression. However, this was later shown to be a mistake because EGFR expression, as currently evaluated, did not correlate with clinical response (48,49) or with breast cancer cell line inhibition (50). Clinical trials in breast cancer have clearly demonstrated that EGFR inhibitors and therapeutic antibodies are targeting EGFR receptors in both tumor and skin from patients, so the lack of effect is not due to the drug’s failure to reach the target or the target’s lack of inhibition (41). So EGFR inhibitors hit their target, and certain cancer patients clearly benefit from these drugs. Therefore, it is surprising that although EGFR-targeted therapy has been in focus for the past decade and drugs are now being administered to patients, no FDA-approved test or profile exists at present to guide clinicians on the use of EGFR inhibitors. Studies have shown EGFR gene copy number, mutations, and downstream signalling to be associated with clinical response in a series of lung cancer trials (51–55). However, these were retrospective, poorly powered studies, which are of insufficient impact to influence molecular diagnostics. Taken together, what has emerged from clinical trials in breast and other cancers along with findings from in vitro studies is that EGFR inhibition is not similar to HER2/neu-targeted therapy in breast cancer, where a single, well-characterized molecular event can be used to target treatment. Therefore, what constitutes a good predictive marker for EGFR-targeted therapy in breast cancer is currently not clear. For future clinical trials with EGFR inhibitors to be successful and current, approved EGFR inhibitors to be administered to the right patients, it will be necessary to understand better the biology of EGFR-directed signaling and the relative importance of mutations, gene copy number, and
downstream signaling pathways have on therapeutic response. These findings can then be used to setup adequately designed molecular pathology and translational studies within clinical trials that will address predictive markers for EGFR-targeted therapy.

Predictive Markers in Drug Development

Future clinical trials should utilize findings from HER2/neu and EGFR-targeted therapy. We have learnt that overexpression or amplification alone does not necessarily mean that protein can be targeted with drugs or that it is indicative of aberrant activation. Posttranslational modifications such as phosphorylation, ubiquitination and acetylation, and localization are critical for protein function. Furthermore, what has become clear is that preclinical models do not reflect the complexity of human tumors well, and, therefore, predictive marker selection must be validated in clinical settings. Despite all these challenges, drug development must try and entail predictive marker development so that stratification of patients can be made possible, thus avoiding unnecessary side effects and improving the chance of a trial being successful. Without an increased focus on predictive marker-directed patient selection, many of the drugs currently in clinical trials run the risk of failing to reach the clinic despite being efficacious for a subpopulation of patients. Therefore, it should be cost effective for pharmaceutical companies to build predictive marker strategies in drug development, as this may improve the success rate of drug development in clinicals. In addition, it would be useful to develop predictive markers of response to many of the drugs already in the market, thus improving patient care and reducing treatment costs by only treating patients who benefit from the treatment.

The major risk we face at present is that we will continue to use poorly designed retrospective analyses of existing trial populations to seek to identify markers on the basis of incomplete understanding of the biology of drug response. This is, in part, due to the fact that we have no reference criteria against which to judge biomarker development. For ERα, despite no level-1 evidence being available, it is unthinkable that this biomarker be regarded as anything less than essential to the management of breast cancer in the twenty-first century. For HER2, only after the central importance of this biomarker was established were steps put in place to produce robust, accurate, and clinically viable test methods. However, clearly for future biomarkers, a structured approach will improve the likelihood of selecting appropriate markers. For this to succeed, we need to progress from the current model to one where, as with drug development, critical attention is paid to each step of the process by which novel biomarkers are developed (56).

To develop and validate predictive markers, a number of hurdles have to be passed before they can be considered for clinical practice. These processes should be addressed in a structured manner to maximise the possibility of success. For example, the development of robust and accurate diagnostic assays should
precede, not follow, the identification of a potential clinical role for biomarkers (56). To illustrate this, we have adapted a biomarker development scheme from Pepe et al. and Hayes et al. (57,58) into four phases (Figure 2).

- Phase 1: Discovery stage. In order to discover predictive markers of response, high-throughput gene expression microarrays or proteomic technologies are often used. The primary output of this phase is to develop a set of potential predictive markers that allow selecting patients who respond to treatment. Ideally, samples from the neoadjuvant setting should be used so that information about the influence that treatment has on the markers selected can be obtained.

- Phase 2: Development stage. In this stage, the predictive marker assay now needs make the transition from the research setting to the clinical diagnostic laboratory. At this stage, it is required to develop the assay for correct determination of predictive marker of response to a level where it
can be used by clinical institutions, according to guidelines put forward by ASCO and EORTC (The European Organization for Research and Treatment of Cancer) (58,59). This means that proper instructions need to be in place for all aspects of predictive marker testing, including sample handling and pretreatment, protocols and standards for ensuring uniform testing across institutions, and external quality assurance. It is desirable that a set of positive controls are developed for interlaboratory variation to be accounted for in the same fashion as the HercepTest comes with positive control cell lines.

- **Phase 3: Validation phase.** In order to validate markers found in phase 1, retrospective validation studies need to be set up. When setting up clinical trial at this stage, it is important that one of the endpoints of the trial is to assess predictive markers. One of the primary outcomes of this phase is to evaluate how well the predictive markers perform in independent validation studies and what predictive markers perform compared better with traditional diagnostic markers such as nodal status, grade, etc.

- **Phase 4: Clinical utility phase.** If success is obtained in phase 3, the marker can be move to phase 4, with clinical studies utilizing randomized prospective screening with predictive markers. The primary outcome of phase 4 is to evaluate if patients selected with predictive markers were the ones that benefited from treatment and if it is ethically defendable to select patients on the basis of the results of the clinical studies. The predictive marker then needs to go through regulatory approval for the defined indication before being applied in the clinic.

Predictive marker development requires the joint effort of multiple academic institutions and industry to cover all the aspects of development. It takes many years and large investments, and the risk for failure is high because of the nature of drug development. That may be the reason why so few new markers have been introduced in recent years despite major research and development efforts and new techniques.

**Future Perspectives**

Cancer is a disease that involves the complex interplay of multiple genes in a framework of mutations in the cancer cell. In addition, tumor cells have interactions with other cells, thereby further promoting growth and invasiveness. Patients respond differently to treatment based on their tumor characteristics. Despite recent challenges, we foresee the strong likelihood that current research and drug development will introduce more predictive markers into IHC laboratories in the same fashion as trastuzumab has resulted in HER2/neu being routinely tested for (see above). This foresight is based on the premise that more care and more robust approaches are applied to the development of predictive biomarkers. It is likely that within two to five years, a panel of markers will
be tested on every tumor sample in order to decide optimal therapy regimen. Currently only three predictive IHC markers (ERα, PR, and HER2/neu) are tested for routinely in breast, but already some pathologists use additional markers to inform decisions, e.g., p53, Ki67, and TOP2A. As more and more predictive IHC markers are introduced in clinical laboratories, this will lead to an increased pressure on pathology laboratories to deliver fast and accurate decisions. How can this be achieved and what technological advances can improve IHC testing?

Current IHC testing in pathology laboratories is subjective and semi-quantitative, as slides are scored manually and using chromogenic dyes (56–61). The use of computer-assisted analysis of predictive markers has been proven to reduce slide-scoring variability among pathologists (31). While FISH/CISH analysis is more robust, there remains room for improvement in this area also. We foresee the introduction of image analysis as a standard tool for IHC scoring, which will begin the process of moving the analytical phase from the microscope to the computer screen.

As analytical challenges increase, accurate quantitation will become increasingly important. Already evidence suggests that additional information could be derived if it were possible to robustly quantify ERα and Ki67. There is, therefore, a strong opportunity for systems such as the Automated QUantitative Analysis (AQUA™) from HistoRX, Inc., New Haven, Connecticut, U.S., to pave the way for a switch from traditional chromogenic staining techniques to fluorescent staining techniques which are potentially quantitative (wide dynamic range) (62). A further advantage is that fluorescent labels can be multiplexed together, yielding spatial information and quantization of proteins in different compartments of the cell. New techniques being developed, such as the proximity ligation in situ assay, show promising results for quantitative analysis of specific protein interactions in FFPE, i.e., formalin-fixed, paraffin-embedded, tissue (63). In some cases, this may be superior to the measurement of protein concentrations or posttranslational modifications, as most proteins take part in multiprotein complexes and their inclusion in the complex indicates relevant activity. In addition, several protein-interaction inhibitors are currently in clinical trials, and, therefore, measurement of the amount of target (interacting proteins) in the tissue may be a predictive marker for responsiveness to the drug.

There remains a key question: Will we be testing protein and DNA markers on slides at all in 5 to 10 years? There are a number of developments in process now miniaturizing assays onto small chips that will measure hundreds of protein or DNA markers quantitatively and in a reproducible fashion. Such technological advances will change the pathology workflow in the future so that every cancer sample coming into the clinic will be analyzed by the pathologist, tumor cells microdissected into solution, and DNA, protein, and RNA extracted and applied onto analytical chips. The readout from these analyses could then be used to determine optimal therapy regimen for the patient and hopefully cure the disease. However, until these techniques have been properly validated and developed for clinical use and trials have proven their value, we are likely to rely
Predictive Markers for Targeted Breast Cancer Treatment

on traditional in situ techniques for protein analysis (IHC) and molecular determination of gene amplifications (FISH/CISH).

In summary, introduction of quality control, automated platforms, and computer-assisted scoring will lead to higher throughput and more accurate decisions with regard to treatment and facilitate the development of future predictive markers.

CONCLUSION

As our understanding of cancer as a disease progresses and the tools to investigate cancer biology develops, the complexity of the disease is unraveling. Breast cancers are now being classified not only on the basis of traditional pathology traits but also according to molecular phenotype and, more recently, gene expression profiles. New molecular targeted therapeutics will lead to more predictive markers being introduced, and these markers must be prospectively validated, internally and externally quality assured, and rigorously quantitative to allow correct decisions for treatment. Therefore, an evolution of the pathology laboratory is likely to take place in the near future, leading to a switch to fluorescent-based techniques, digital imaging of slides, and computerized scoring of predictive markers. On a longer time horizon, new technologies being developed and applied are likely to change cancer diagnostics so that molecular phenotyping of patients will take center stage in patient management. In the not too distant future, there is the possibility that pathologists will type and stage tumors and simultaneously identify key areas of tissue for automated microdissection and processing. The subsequent process of molecular phenotyping, which may include kinomic, proteomic, and transcriptomic assays performed simultaneously could be carried out in specialised laboratories using robotics and microfluidics on high-density protein and gene arrays to determine optimal treatment regimens for tumors with specific molecular profiles linked to drug response. Such a future vision may represent the more radical view of what could be accomplished by linking of current concepts and technologies to maximize patient benefit from therapeutic options. However, we must dream of such futures if we are to maximize progress in the area of molecular pathology for the benefit of the most important people, our patients.

REFERENCES


Predictive Markers for Targeted Breast Cancer Treatment 147


Predictive Markers for Targeted Breast Cancer Treatment


Circulating Tumor Cells in Individualizing Breast Cancer Therapy

James M. Reuben
Department of Hematopathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, U.S.A.

Massimo Cristofanilli
Department of Breast Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, U.S.A.

INTRODUCTION

Breast cancer is the most frequently diagnosed nonskin cancer among women in the United States and is second only to lung cancer in causing cancer-related deaths among them (1). The vast majority of deaths are due to recurrent metastatic disease. Occult dissemination of tumor cells is the main cause of recurrent metastatic breast cancer (MBC) in patients who have undergone resection of their primary tumor (2). Approximately 5% of patients with breast cancer have clinically detectable metastases at the time of initial diagnosis, and a further 30% to 40% of patients who appear clinically free of metastases harbor occult metastases (3,4). The formation of metastatic colonies is a continuous process, commencing early during the growth of the primary tumor. Metastasis occurs through a cascade of linked sequential steps involving multiple host-tumor interactions. This complex process requires the cells to enter the circulation, arrest at the distant vascular bed, extravasate into the organ interstitium and parenchyma, and proliferate as a secondary colony.
Several experimental studies suggest that during each stage of the process, only the fittest tumor cells survive (2).

The first report on tumor cells in the peripheral circulation was attributed to Ashworth in 1869 (5). Since then, the existence, origin, and clinical significance of circulating tumor cells (CTCs) have been debated. The introduction of sensitive and specific immunohistochemical techniques in the late 1970s led to renewed interest in the detection of CTCs and their possible association with minimal residual disease in solid malignancies (6,7). However, there is little understanding of the genetic and phenotypic characteristics of CTCs in peripheral blood. It has been speculated that mammary stem cells represent the cellular origins of cancer because they exist quiescently over long periods and can accumulate multiple mutations over the lifetime, ultimately giving rise to tumors when stimulated to proliferate (8–10). It was recently reported that highly tumorigenic cells possessing properties consistent with those of stem or progenitor cells could be isolated from human breast cancers (9). Furthermore, transplantation of as many as several hundred of these cells into the cleared mammary fat pads of etoposide-treated mice resulted in the growth of breast tumors (10). Cancer stem cells not only exist in human breast cancer but also have different biologic features than the more differentiated cells that constitute the majority of cells in human breast cancers. These cancer stem cells may be detectable in the peripheral blood and bone marrow of patients with MBC and may account for their worse prognosis. Hence the development of therapies specifically targeting cancer stem cells may have great benefits for patients with CTCs who possess the properties of cancer stem cells.

When CTCs migrate to distal sites, they can potentially give rise to microscopic lesions that result in metastasis. Microscopic disease may be present in lymph nodes, bone marrow (primary breast cancer), and peripheral blood (metastatic disease) at the time of breast cancer diagnosis (3,4,11,12). Most studies have demonstrated that the detection of microscopic disease in patients with breast cancer contributes prognostic information and, in selected cases, can predict the efficacy of treatments (12). In patients with primary breast cancer, the detection of microscopic disease in lymph nodes and bone marrow has led to a better understanding of the role of minimal residual disease in establishing metastasis.

Over the past few years, immunomagnetic separation technology, with its higher level of sensitivity and specificity, has been used to improve the detection of CTCs compared with the detection of occult CTCs by reverse transcriptase-polymerase chain reaction (RT-PCR) (13–21). In this chapter, we review the methods of detecting CTCs, the prognostic implications of CTCs, and the potential means of exploiting CTCs in developing individualized therapy.

DETECTION OF CIRCULATING TUMOR CELLS

Advances in technology have facilitated the detection of even very small numbers of CTCs in the peripheral blood of cancer patients. The principle of these methods is based on the fact that breast cancer cells express epithelial cell
adhesion molecule-1 (EpCAM). EpCAM, a 40-kDa glycoprotein, is present on most epithelial carcinomas, and recent studies indicate that this molecule has a major morphoregulatory function, relevant not only to epithelial tissue development but also to carcinogenesis and tumor progression (22,23).

Detection of CTCs by the CellSearch™ System

The only assay approved by the U.S. Food and Drug Administration is the CellSearch™ system (Veridex, LLC, Warren, New Jersey, U.S.). With this assay, a small sample of peripheral blood is allowed to react with ferrofluids coated with antibodies to EpCAM. Next, the sample is processed through a magnetic field to retain the EpCAM-positive cells, while the EpCAM-negative cells, primarily of hematopoietic origin, are eluted from the column and discarded. Thereafter, the enriched EpCAM-positive cells are allowed to react with the nucleic acid dye 4′,6-diamidino-2-phenylindole (DAPI), monoclonal antibodies specific for leukocytes (anti-CD45 antibodies conjugated with allophycocyanin), and epithelial cells (phycoerythrin-conjugated antibodies to cytokeratin 8, 18, and 19) and analyzed by the CellSpotter™ Analyzer (Veridex). The CellSpotter Analyzer is a semiautomated fluorescence-based microscopy system that enables computer-generated reconstruction of cellular images. The analyzer interrogates each cell image to determine if it meets the very stringent requirements of an algorithm to be classified as a CTC. The algorithm insures that a CTC expresses EpCAM and not leukocyte lineage-specific antigens (represented by CD45), exhibits cytoplasmic expression of cytokeratin, and contains a nucleus that binds DAPI. Absence of any of these characteristics disqualifies a cell image as a CTC. In identifying CTCs, the procedure identifies cell objects that possess some, but not all, of the required characteristics to be a CTC and are labeled as “unclassified” objects (24). The stringent criteria of CellSearch for identifying a CTC are laudable; however, the merits of the assay continue to be debated as it minimizes the scientific or clinical significance of unclassified objects.

Although CellSearch is a very reproducible and reliable method for the identification and enumeration of CTCs in the peripheral blood of patients with breast cancer, its clinical utility has been limited by its high cost. The semiautomated assay is also somewhat labor intensive. In addition, the need to permeabilize the cell membrane of a viable cell to introduce DAPI and anticytokeratin antibodies to label intracellular structures further deters the possibility of interrogating viable cells for their growth potential, colony formation, and genomics, among others. To address these issues, the CellProfile™ kit was introduced that permits the retrieval of EpCAM-positive cells prior to cell permeabilization and interrogation for genomic or proteomic profiles of the CTCs.

Isolation of CTCs by the AutoMACSTM System

EpCAM, also known as human epithelial cell antigen (HEA) or CD326, is expressed on the majority of tumor cells of epithelial cell origin but not on
circulating B cells, T cells, or monocytes. In another assay that exploits this characteristic of CTCs, peripheral blood mononuclear cells are incubated with magnetic beads coated with anti-CD326 (Miltenyi Biotec, Inc., Auburn, California, U.S.) prior to passage through a magnetic column to enrich for CTCs. Typically, mononuclear cells are isolated from peripheral blood, bone marrow, pleural fluid, paracentesis fluid, or other tissues, mixed with anti-CD326 microbeads, and incubated at 4°C to 8°C for 15 minutes. After washing, the cell pellet is resuspended and loaded onto the magnetic column of an automatic magnetic cell sorting (AutoMACSTM) system (Miltenyi Biotec) (25). CD326-positive CTCs are isolated using the positive selection protocol and can then be centrifuged onto glass slides to determine the morphology, viability, and purity of the preparation.

In one study, Papanicolaou (Pap)-stained cytospin smears were analyzed in detail to identify tumor cells on the basis of conventional cytologic criteria. In cytospin slides of enriched CTCs, large cells with cytomorphological features suspicious for tumor cells were detected, and subsequently the cytospin slides were immunocytochemically stained with the avidin-biotin complex peroxidase technique using pancytokeratin antibody for the identification of tumor cells (Fig. 1A). (The immunostains of CTCs were performed and interpreted by Dr. Savitri Krishnamurthy, Department of Pathology, The University of Texas M. D. Anderson Cancer Center).

A second cytospin slide of CD326-positive bone marrow tumor cells from the same patient was stained with a cocktail of anticytokeratin antibodies using the avidin-biotin complex peroxidase technique for the identification of tumor cells (Fig. 1B). In the majority of cases, the suspicious cells were positive for cytokeratin, thereby confirming the presence of micrometastatic carcinoma.

In addition to bone marrow, it is possible to isolate CTCs from the paracentesis fluid of a patient with MBC. Figure 2 shows a cluster of CTCs

**Figure 1A** A cytospin slide of CTCs from the bone marrow of a patient with primary breast cancer using anti-CD326 antibody–coated microbeads and reacted with Pap stain. The figure shows a cluster of CTCs with cellular morphology consistent with that of epithelial cells. Abbreviations: CTCs, circulating tumor cells; Pap, Papanicolaou.
Isolation of CTCs by the EasySep™ System

Another immunomagnetic cell selection system is EasySep™ (StemCell Technologies, Inc., Vancouver, Canada), which combines the specificity of monoclonal antibodies with the simplicity of a column-free magnetic system. In this procedure, highly enriched cells are obtained by positively selecting the cells of interest with antibodies to a specific cell-surface antigen. The cells targeted for isolation from the paracentesis fluid of a patient with MBC that were reacted with Pap stain.

Figure 1B  CD326-positive CTCs are cytokeratin-positive epithelial cells. Abbreviation: CTCs, circulating tumor cells.

Figure 2  CD326-positive CTCs from the paracentesis fluid of a patient with MBC. Abbreviations: CTCs, circulating tumor cells; MBC, metastatic breast cancer.
selection are cross-linked to EasySep nanoparticles via the formation of tetrameric antibody complexes (StemCell Technologies) in a standard centrifuge tube, which is then placed in the EasySep magnetic chamber. The handheld magnetic chamber is gently inverted to discard the cells that are not bound to the specific nanoparticle complexes that adhere to the walls of the tube in the magnetic chamber. To obtain a highly pure population, the residual cells in the tube are rinsed twice and then harvested by removing the tube from the chamber. Enriched CTCs can then be used for further interrogation for the differential expression of EpCAM or other markers of interest.

Positive selection is most effective when a specimen has a generous complement of the desired cell population. When a specimen, such as peripheral blood, contains few or inadequate numbers of the cells of interest, a negative-selection approach can be taken to enrich for these cells. One negative-selection technique is the RosetteSep™ (StemCell Technologies) method that uses a highly purified combination of mouse and rat monoclonal antibodies. When reacted with peripheral whole blood, these antibodies bind in bispecific tetrameric antibody complexes directed against cell surface antigens on human hematopoietic cells (CD2, CD16, CD19, CD36, CD38, CD45, and CD66b) and glycophorin A on red blood cells, or erythrocytes. The unwanted cells of hematopoietic origin form rosettes by cross-linking to multiple erythrocytes that can be easily removed by performing a typical Ficoll-Paque™ (GE Amersham-Pharmacia Diagnostics, Uppsala, Sweden) density gradient procedure for the separation of mononuclear cells from peripheral blood. The rosettes, free erythrocytes, and granulocytes form a pellet, while the unlabeled, desired CTCs are collected from the interface between the plasma and the buoyant density medium. Figures 3A and 3B represent cytospin slides of CTCs isolated from the peripheral blood of a patient with MBC using RosetteSep™.

**Figure 3A** Pap stain of a cluster of CTCs isolated from the peripheral blood of a patient with MBC using RosetteSep™. *Abbreviations:* Pap, Papanicolaou; CTCs, circulating tumor cells; MBC, metastatic breast cancer.
peripheral blood of a patient with MBC that were reacted with Pap stain (Fig. 3A) and anticytokeratin antibodies (Fig. 3A).

Detection of CTCs by AdnaGen's Technology

In addition to magnetic separation methodologies, PCR greatly facilitates the detection of occult tumor cells through the use of nucleic acid analysis. The PCR-based assay, Adna Test Breast Cancer Select, developed by AdnaGen AG, Langenhagen, Germany, makes use of RT-PCR to identify putative transcripts of genes in EpCAM-positive cells that are isolated by a magnetic separation method. This approach takes advantage of the prevailing state of the art of CTC isolation and enhances the chance of detecting very low numbers of CTCs or occult CTCs on the basis of their expression of tumor-associated genes. AdnaGen’s two-step “combination-of-combinations principle” involves initially the enrichment of CTCs using an antibody mix linked to magnetic particles (26). Thereafter, mRNA is extracted from the CTCs and transcribed into cDNA before being subjected to multiplex RT-PCR for the detection of tumor-associated genes. Owing to the combination of different selection and tumor markers, both the heterogeneity of the tumor cells and the possible therapy-induced deviations in the expression patterns are taken into account. The AdnaGen system has high analytical sensitivity (two cells in 5 mL of blood) and high specificity (>90%) achieved through the combination of multimarker tumor cell enrichment and multiplex gene expression profiling.
CLINICAL SIGNIFICANCE AND PROGNOSTIC VALUE OF CTCs in MBC

Tumor cells can be detected in the locoregional lymph nodes of women with early breast cancer, and the presence of these cells has been shown to have a negative effect on long-term prognosis (3, 12). Despite evidence of the prognostic value of CTCs in some studies, the detection of micrometastases was never incorporated into cancer-staging protocols or considered a valuable clinical tool. This may be the result of a combination of factors, such as variable antigen expression in poorly differentiated tumors and reports of cytokeratin and epithelial membrane antigen positivity in cells that are not of epithelial origin, which demonstrated a need for more sensitive and specific methods of detection than were available at the time. This need was filled to some degree by sensitive PCR techniques (14, 16, 17). In the past decade, a few studies have shown that the detection of occult disease by PCR has prognostic significance in some solid tumors (13). However, PCR-based assays for the detection of occult tumor cells have limitations, particularly contamination of samples, specificity of the assays, and inability to quantify tumor cells. Moreover, PCR-based methods cannot be used to perform functional assays. These factors have precluded the widespread use of PCR in in vitro diagnostic applications.

Using an immunomagnetic detection approach, Austrup et al. reported the prognostic significance of genomic alterations (e.g., c-erbB-2 overexpression) present in CTCs purified from the blood of patients with breast cancer (27). The authors investigated genomic imbalances, such as mutation, amplification, and loss of heterozygosity, of 13 tumor suppressor genes and 2 proto-oncogenes using DNA isolated from minimal residual cancer cells. The presence and the number of genomic imbalances measured in disseminated tumor cells were significantly associated with a worse prognosis (27).

Subsequently, Meng et al. demonstrated that CTCs recapitulate the human epidermal growth factor receptor 2 (HER2) status of the primary tumor (28). Furthermore, the authors demonstrated that a fraction of patients with HER2-negative primary tumors had detectable HER2 gene amplification in their CTCs, suggesting acquisition or selection of this phenotype during cancer progression (37.5%; 95% confidence interval, 18.8–59.4%). Intriguingly, the initiation of trastuzumab-based therapy in a few cases with altered HER2 status in CTCs was associated with a clinical response (29). These important observations support the argument that detection of CTCs and determination of their gene amplification or expression can be used for better tailoring of therapies.

The predictive and prognostic roles of CTCs were investigated in a prospective multicenter clinical trial led by researchers at the University of Texas M. D. Anderson Cancer Center (11). In this study, the CellSearch system was used to prospectively determine the prognostic and predictive value of CTCs in patients with MBC who were about to start a new systemic treatment. The 177 enrolled patients underwent peripheral blood collection at monthly intervals for up to six months after enrollment. A cutoff of 5 CTCs/7.5 mL was used to
stratify patients into positive and negative groups (positive, ≥5 CTCs/7.5 mL; negative, <5 CTCs/7.5 mL). The investigators reported that patients classified as positive had shorter progression-free survival times (2.7 months vs. 7.0 months; \( P = 0.0001 \)) and shorter overall survival times (10.9 months vs. 21.9 months; \( P < 0.0001 \)) than did those classified as negative. Furthermore, the CTC status at first follow-up after initiation of therapy (three weeks) had an even greater association with progression-free survival time (2.1 months versus 7.0 months; \( P < 0.0001 \)) and overall survival time (8.2 months versus >18 months; \( P < 0.0001 \)). On multivariate Cox hazards regression analysis, CTC levels, both at baseline and at first follow-up, were the most significant predictors of progression-free and overall survival (11).

An analysis restricted to patients who were about to start first-line systemic therapy after the diagnosis of MBC had similar findings (30). A subsequent investigation including patients with either measurable or evaluable MBC confirmed that CTCs are an independent predictor of survival that is more powerful than standard prognostic measures, such as hormone receptor status, and measures of tumor burden (e.g., the Swenerton score or level of CA27-29) (31).

The detection of changes in CTC status may also have predictive utility. The CTC detection rate at first follow-up was lower than at baseline, particularly in patients undergoing first-line treatment (25% versus 52%) and those with visceral disease (28% versus 50%) (11). In patients with newly diagnosed disease treated with chemotherapy alone (n = 37), the percentage of patients who were CTC positive decreased significantly from baseline to first follow-up (57% to 37%, \( P = 0.001 \)) and first imaging-visit (at eight to nine weeks) blood draw (57% to 31%, \( P = 0.004 \)) (31). More importantly, in patients receiving trastuzumab-containing regimens, CTC-positive patients decreased from baseline to first follow-up (59% to 7%, \( P = 0.600 \)) and first imaging-visit blood draw (59% to 0%, \( P = \text{not applicable} \)). This data indicates that patients with newly diagnosed disease who are about to start first-line therapy can have detectable CTCs and that the changes in CTC status at three to four weeks may indicate a benefit from systemic treatment (particularly trastuzumab-based regimens). Moreover, the determination of CTCs at baseline and follow-up appears to be a superior prognostic marker in patients with measurable MBC compared with standard imaging assessments (32). CTCs could be used for assessment of the clinical benefit of systemic treatments, for prognostic stratification, and for evaluation of patients with nonmeasurable disease in future prospective validation trials.

STAGING MBC USING CTCs

The American Joint Committee on Cancer (AJCC) classification system, based on a schema developed by the International Union Against Cancer (UICC) (33–35), includes much of the traditional prognostic information used by clinicians when developing a comprehensive treatment plan. In brief, the AJCC
system attempts to define the disease by incorporating all aspects of cancer distribution in terms of the primary tumor (T), lymph nodes (N), and distant metastasis (M). A “stage” group (0, I, II, III, or IV) is then assigned on the basis of the possible TNM permutations, with 0 reflecting minimal involvement and IV either the maximum tumor involvement or distant metastasis. This basic TNM staging is then further subdivided according to the time the evaluation is performed: c = clinical (before surgical treatment) and p = pathologic (after surgical specimen analysis) (10,35).

Over the past 45 years, the AJCC has regularly updated its staging standards to incorporate advances in prognostic technology. The committee’s current work concerns the development of prognostic indices based on molecular markers. However, until these changes are incorporated, TNM staging continues to quantify only the physical extent of the disease. Although it covers the approximately 2% of women who present with metastatic disease (stage IV), the prognostic information is applicable only to in situ, local, and regional primary breast cancers. Given the heterogeneity of the disease in primary and metastatic settings, the potential for continued mutation, and the variety of treatment options available, the information acquired at the time of the initial diagnosis of breast cancer may not be as relevant to planning the treatment for recurrence in the approximately 30% of women who have recurrent MBC years after their initial diagnosis.

The recent demonstration that the presence of CTCs predicts the prognosis of two subgroups of patients with MBC raises the possibility that this method will allow for a true “biologic staging” of breast cancer (e.g., stages IVA and IVB) (11). The main limitations of the previous study were the sample size (only 177 patients); the inclusion only of patients with measurable disease, which does not reflect the heterogeneity of MBC; and the inclusion of patients at different points in treatment—those with newly diagnosed disease and those undergoing second- or third-line treatment. To overcome these limitations, a larger international validation trial, the International Stage IV Stratification Study, was begun. The objective of this study was to confirm the prognostic value of CTC detection in a larger and more homogeneous cohort of patients with newly diagnosed MBC. The study aimed at enrolling 660 patients and providing a more detailed analysis of the association between CTC detection and other factors such as ethnicity (black compared with white and Hispanic groups) and specific disease sites (visceral versus nonvisceral disease).

**CLINICAL APPLICATION OF CTCs**

Despite years of clinical research, the odds of patients with MBC achieving a complete response remain extremely low. Thus, the main goal in the management of MBC is palliation (36). Only a few patients who experience a complete response after chemotherapy remain in this state for prolonged periods, although
some have remained in remission for more than 20 years (37). These long-term survivors are usually young, have an excellent performance status, and, more importantly, have limited metastatic disease (37,38). Most patients with MBC respond only transiently to conventional therapies and develop evidence of progressive disease within 12 to 24 months of starting treatment. For these patients, systemic treatment does not result in a significant improvement in survival time but may improve their quality of life.

At present, clinicians use three different systemic treatment modalities for advanced breast cancer: endocrine therapy, chemotherapy, and biologic targeted therapy (39–42). Appropriate selection of patients for these modalities is based mostly on tissue assessment of hormone receptor status (estrogen and progesterone receptors) and c-erbB-2 status. In patients who lack expression of hormone receptors or who demonstrate no amplification of c-erbB-2, cytotoxic chemotherapy is used. However, no standard therapeutic regimen has been defined. For example, the optimal schedule of chemotherapy administration in MBC (i.e., concurrent vs. sequential) remains controversial, and the decision must be individualized. Sledge et al. addressed this issue in a prospective study that included 739 chemotherapy-naïve patients who were randomly assigned to receive, at progression, doxorubicin, paclitaxel, or both (39). Although the response rates and times to treatment failure were improved with the combination regimen, the overall survival rate was comparable in all groups. Other trials have demonstrated a survival advantage with combination regimens over single-agent chemotherapy, but all of these studies have shown differences in the actual median overall survival time between 10 and 13 months across studies, suggesting that even with comparable inclusion criteria, the heterogeneity that is typical of MBC cannot be eliminated (41,42). This heterogeneity is indeed one of the major limitations to the development of more personalized treatments for patients. Therefore, although the data demonstrate that patients can benefit from combination therapy, they do not clearly identify the subsets of patients who will most benefit or who will experience only additional toxicity.

In this context, the use of CTCs to stratify patients at the time of disease recurrence may be an appropriate way to design personalized therapeutic approaches. CTC detection may enable a more rational selection of treatments for patients with newly recurrent disease, and this approach could maximize the chance of a particular combination or single new drug showing clinical benefit and, eventually, prolonging survival. It is possible that CTC detection could be used in the design of efficacy trials of different therapeutic approaches. The efficacy of these treatments could be more easily assessed if patients were stratified by their prognosis, leading to more tailored treatment strategies. We believe that the challenge for the next generation of clinical trials, and the responsibility for both clinical investigators and the pharmaceutical industry will be to incorporate these concepts into the process of drug development.
FUTURE USES OF CTCs

The process of sorting cancer cells from other cellular components (e.g., blood and stromal cells) in clinical samples is fundamentally important for the future of genomic and proteomic analysis (43–45). Collecting representative tissue from solid tumor metastases usually requires more invasive procedures that increase the risk of complications and discomfort. Furthermore, these procedures may not provide an adequate specimen for detailed analysis and typically cannot be repeated for dynamic evaluation of the biologic changes during treatments. Theoretically, CTC detection would allow specific genes [e.g., c-erbB-2, epidermal growth factor receptor (EGFR), and mammaglobin B (MGB)] or more global gene expression to be analyzed while using specific targeted treatments for MBC based on the expression of the CTCs (29,46–48). This information could then be used to design specific treatments that more appropriately reflect the dynamics and heterogeneity of MBC.

CONCLUSIONS

The detection of microscopic disease in the peripheral blood of patients with MBC provides prognostic information. This information will allow appropriate risk stratification and modification of the current staging system for advanced disease. Furthermore, the intriguing hypothesis that CTCs are, at least in part, representative of tumorigenic cancer stem cells may present an additional challenge for translational research. In fact, the phenotypic evaluation of CTCs is indicating the possibility of more critical design of targeted therapies in the setting of MBC. Moreover, with the introduction of this novel concept, it is expected that investigators and the pharmaceutical industry will derive benefits from these sorting technologies that will contribute to more tailored treatments and sophisticated trial designs.

REFERENCES

5. Ashworth T. A case of cancer in which cells similar to those in the tumours were seen in the blood after death. Aust Med J 1869; 14:146.


INTRODUCTION

Discovery of estrogen receptors (ER) was critical for the development of endocrine therapy in breast cancer. Expression of ERα, the predominant isoform, in breast tumors of both premenopausal and postmenopausal women is a highly predictive marker for response to antiestrogen treatment in women with ERα-positive breast cancer. Today, endocrine therapies include the use of antiestrogens such as tamoxifen and aromatase inhibitors such as anastrozole, exemestane, and letrozole. Tamoxifen, the first antiestrogen to be used for the treatment of ERα-positive breast cancer, competitively blocks the actions of 17β-estradiol (E2), the female hormone that binds and activates ERα in tumors. In postmenopausal women, peripheral aromatization of androgens to estrogens is the major source of plasma estrogen. Aromatase inhibitors inhibit this reaction and consequently suppress the production of circulating estrogen in postmenopausal women. Endocrine therapies are effective at reducing recurrence, increasing overall survival, and reducing contralateral breast cancer up to 50%. However, about 50% of patients with ERα-positive breast cancer have
intrinsic resistance to antiestrogen therapy and therefore do not benefit. In contrast to patients with intrinsically resistant tumors, there are patients who do initially respond to antiestrogen therapy; however, most of these patients develop acquired resistance during the treatment regimen. Therefore, the current goal in breast cancer research is to elucidate the mechanisms of both intrinsic and acquired resistance to tamoxifen and the aromatase inhibitors in order to develop new therapeutic strategies to prevent and/or treat resistant breast cancer.

ANTIESTROGENS

Nonsteroidal antiestrogens were initially developed as contraceptives in the 1960s. Walpole and colleagues synthesized tamoxifen (termed ICI 46, 474), a potent antiestrogen with antifertility properties in rats. However, in humans, tamoxifen induced ovulation in subfertile women. Therefore, the development of tamoxifen as a contraceptive was terminated. However, Walpole also patented the application of tamoxifen as a drug treatment for hormone-dependent cancers. Thus, clinical trials were started to evaluate tamoxifen against the standard endocrine treatment at the time, diethylstilbestrol (DES), for the treatment of advanced breast cancer in postmenopausal women (1). Tamoxifen not only was as effective as DES for the treatment of advanced breast cancer but also had fewer side effects. Therefore, the advantage of tamoxifen over DES was crucial for its subsequent evaluation as a treatment for all stages of breast cancer.

In 1962, Jensen and colleagues discovered ERα. Jensen demonstrated that E2, the circulating female hormone that promoted breast cancer growth, binds to diverse tissue sites around a woman’s body but is retained in estrogen-target tissues, for example, the uterus and vagina (2). The identification of ERα as the target of E2 action in the breast and antiestrogens blocking the binding of E2 to ERα provided a therapeutic target and an approach for the treatment of breast cancer. Although many antiestrogens were discovered and tested during the 1960s and 1970s, only tamoxifen was considered safe enough for extensive clinical evaluation (3). Clinical trials ultimately demonstrated that patients with ERα-positive breast cancer benefited the most from tamoxifen therapy, whereas women with ERα-negative breast cancer were found to be unaffected. During subsequent clinical trials, five years of adjuvant tamoxifen treatment was found to be more effective than less than five years of treatment in improving time to tumor recurrence and overall survival (4). In contrast to the beneficial effects of tamoxifen as a treatment for breast cancer, both laboratory and clinical results showed that tamoxifen increased the risk of endometrial cancer in the uterus fourfold in postmenopausal women compared with untreated women (5,6). These results strongly indicated that tamoxifen was not a pure antiestrogen but had selective functions depending on the target tissue.
Resistance to Tamoxifen

Overview

Five years of adjuvant tamoxifen therapy is a standard of care for early-stage ERα-positive breast cancer. Although there is increasing use of aromatase inhibitors in postmenopausal women, tamoxifen is the only approved effective therapy in premenopausal women. Approximately 50% of patients with ERα-positive breast cancer benefit from tamoxifen treatment in the advanced- and early-stage settings (4,7). Unfortunately, the other 50% of patients do not respond, and the ones that do respond initially eventually progress. Thus, the problem with tamoxifen therapy is that breast tumors can either be resistant (intrinsic or de novo) prior to endocrine treatment or become resistant (acquired) during therapy. Numerous studies have looked at the mechanisms responsible for this resistant phenotype. They include the loss of the ERα gene and/or protein in tumors during short-term and long-term treatment, identification of mutations within the ERα gene, changes in the pharmacologic response of resistant breast cancer cells to tamoxifen, activation of ERα-mediated gene transcription in the absence of ligand, modifications in gene transcription by changes in interactions between ERα and coregulators, and the development of enhanced surface-to-intracellular signaling cross-talk between growth factor receptors and ERα (Fig. 1).

Figure 1  Mechanisms of hormone resistance.
ER-Negative Breast Cancer and Resistance

Breast cancers initially lacking expression of ERα have intrinsic resistance to endocrine therapy. In addition, loss of ERα expression upon recurrence of breast cancer during tamoxifen therapy has been reported in as many as 25% of tumors (8–10). However, most of the acquired resistance during adjuvant tamoxifen treatment in ERα-positive tumors is not due to changes in ERα expression. Interestingly, most tamoxifen-resistant breast tumors retain ERα and respond to secondary endocrine treatments. Overall, a loss of ERα status does not seem to be the major mechanism for the development of acquired antiestrogen resistance.

ER Mutations

Several mutant forms of both the ERα and ERβ, another ER isoform, have been identified and previously reviewed (11–13). However, the functional significance of these mutants as a mechanism of antiestrogen resistance is yet to be elucidated. Most of the data demonstrate the existence of splice variants for ERα and ERβ mRNAs without evidence of translation into functional proteins. Most of the tumors express both mutant and wild-type ER with the wild type being the predominant species. A single-point mutation in the ligand-binding domain (LBD) of ERα has been reported (D351Y) that converts tamoxifen and raloxifene from antiestrogen to estrogens in vivo and in vitro models of antiestrogen-stimulated MCF-7 tumor cells (14,15). However, the D351Y ERα mutant has yet to be detected in either intrinsic or acquired resistant breast tumors from patients. In addition, mutations in the F-region of the ER have been shown to affect the activities of both E2 and 4-hydroxytamoxifen (4OHT), the active metabolite of tamoxifen (16). Therefore, the clinical relevance of ER mutants is unclear to date as a mechanism of resistance to antiestrogens. However, our understanding of ER mutations in response to antiestrogens might lead to the development of better antiestrogens or other drugs for breast cancer treatment.

Role of Pharmacogenetics

Tamoxifen undergoes extensive primary and secondary metabolism. The secondary metabolite 4-hydroxy-N-desmethyl tamoxifen (endoxifen) is formed by the CYP2D6-mediated oxidation of N-desmethyl tamoxifen and has been shown to play an important role in the anticancer effect of tamoxifen. CYP2D6 genotyping has revealed that women can be classified as either poor, intermediate, or extensive metabolizers of tamoxifen. Women who are poor metabolizers of tamoxifen, and, therefore, have lower levels of endoxifen, have been shown to have worse disease-free survival. No difference in overall survival has been shown (17). Women who are extensive metabolizers of tamoxifen have also been shown to experience more hot flashes. A recent study showed that women who experience hot flashes while on tamoxifen had a hazard ratio of recurrence of
Individualization of Endocrine Therapy in Breast Cancer

0.51 compared with women without hot flashes (18). This study suggests that vasomotor symptoms and efficacy of tamoxifen may be related to a women’s pharmacogenetics.

Role of Coregulators in Resistance

The recruitment of coactivators or corepressors to the ER determines the switch between ER activation and repression. In addition, the coordinated action of ligand, ER, and coregulators determines which genes are transcribed or repressed depending on the cellular context and thus which cells will or will not proliferate. The overall data indicate that intricate modulation of the ER-to-coregulator ratio in breast cancer cells could determine resistance to antiestrogens. For example, the coactivator SRC-1 activates ER in a ligand-independent manner while increasing 4OHT’s agonist activity (19). On the other hand, the corepressor SMRT blocks the agonist activity of 4OHT-induced by SRC-1 (20). The other member of the p160 coactivator family is amplified in breast cancer (AIB1, SRC-3, and RAC3). AIB1 mRNA was found to be amplified in 60% of breast cancers, and the protein product was found to be overexpressed in about 10% of tumors. Recently, an association was discovered between high AIB1 and human epidermal growth factor receptor 2 (HER2/neu), a member of the epidermal growth factor receptor (EGFR) family of receptor tyrosine kinases and tamoxifen-resistant breast cancers (20). The potential mechanism for this resistance requires activation of the mitogen-activated protein kinase (MAPK) signaling cascade by HER2/neu, which in turn potentially phosphorylates and activates AIB1. Activated AIB1 in turn can convert 4OHT from an antiestrogen to an estrogen in regards to ERα activity. In contrast, low levels of N-CoR mRNA, a corepressor, have been implicated with resistance to tamoxifen (21). Therefore, the identification of abnormally expressed coregulators will probably assist in the design of future therapies that improve the efficacy of endocrine treatment without the development of resistance.

EGFR and HER2/Neu and Antiestrogen Resistance

Numerous laboratory and clinical studies indicate that overexpression and/or aberrant activity of the HER2/neu (erbB2) signaling pathway is associated with antiestrogen resistance in breast cancer (22). The HER2/neu receptor is a member of the EGFR family of receptor tyrosine kinases, which include HER3 (erbB3) and HER4 (erbB4) (23–26). Upon dimerization, the tyrosine kinase domains located within the COOH-terminal regions of receptors are activated by an autophosphorylation cascade on specific tyrosine residues, which activate downstream effectors, such as MAPK and Akt, which promote cellular proliferation, survival, anti-apoptosis, and transformation. HER2/neu is overexpressed and/or amplified in 25% to 30% of breast tumors and is associated with a more aggressive phenotype and poor prognosis (27). Patients with breast tumors overexpressing HER2/neu exhibit much lower response rates to antiestrogen
therapy (27). Thus, it is suggested that one possible mechanism of resistance to tamoxifen is overexpression of HER2/neu in ERα-positive breast cancers. In addition, overexpression of EGFR and its ligands is observed in several human cancers including breast cancer (28–30). The increased expression of EGFR is frequently associated with tumor progression and resistance to antiestrogens. These data suggest that EGFR and/or HER2/neu are possible targets for preventing or treating antiestrogen resistance in breast cancer. Strategies to target EGFR and HER2/neu include the use of humanized monoclonal antibodies to the receptors (31), tyrosine kinase inhibitors that block reduction of adenosine triphosphate (ATP) to adenosine diphosphohate (ADP) + Pi, and receptor antisense molecules (32). Gefitinib, (ZD 1839, Iressa\textsuperscript{1}), an EGFR-specific tyrosine kinase inhibitor, has been shown to inhibit growth of breast cancer cell lines in vitro that are resistant to tamoxifen (33). More importantly, the combination of gefitinib and tamoxifen was shown to prevent resistance (34), demonstrating that (1) EGFR might be a key player in the development of antiestrogen resistance and (2) inhibiting the activity of this receptor might be therapeutically beneficial in preventing resistance to antiestrogens. In addition to gefitinib, trastuzumab (Herceptin\textsuperscript{1}) is a humanized monoclonal antibody directed against the ectodomain of the HER2/neu receptor. It has been shown to restore breast cancer cell sensitivity to tamoxifen in HER2/neu-overexpressing cells (35).

Role of MAPK

MAPK is a serine/threonine kinase that is activated by phosphorylation cascades originating from GTP-bound Ras, a downstream effector of EGFR and HER2/neu. MAPK has been shown to be hyperactivated in MCF-7 breast cancer cells overexpressing either EGFR or HER2/neu (36–38). This increased activity of MAPK promoted an enhanced association of ERα with coactivators and decreased interaction with corepressors, thereby enhancing hormone-dependent gene transcription and possibly leading to tamoxifen-resistant breast cancer. The role of the MAPK signaling pathway in tamoxifen resistance has been demonstrated primarily in vitro using specific inhibitors such as U0126 (MAPKK, MEK1/2 inhibitor), AG1478 (EGFR and HER2/neu inhibitors), and PD98059 (MAPK, ERK1/2 inhibitor) (36–38). The exact mechanism by which hyperactivated MAPK converts the antiestrogenic activity of the tamoxifen-ERα complex to more of an estrogenic one in resistant breast cancer cell is yet unclear. However, it has been demonstrated that activation of the Ras/MAPK pathway by EGFR/HER2/neu receptor activation could phosphorylate Ser-118 in the AF-1 domain of ERα, thus promoting ligand-independent transcription of ERα and possibly loss of tamoxifen-induced inhibition of ERα-mediated gene transcription. On the basis of these preclinical models, several clinical trials are under way to determine whether treating patients with EGFR inhibitors (i.e., Iressa) and/or MAPK-specific inhibitors (U0126) can effectively treat and/or prevent antiestrogen resistance in breast tumors.
Role of PI3-K/AKT Signaling

Phosphatidylinositol 3-kinase (PI3-K) is a heterodimer complex consisting of a regulatory subunit, p85, and a catalytic subunit, p110. The regulatory subunit p85 is phosphorylated by either receptor tyrosine kinases such as EGFR/HER2/neu or intracellular adapter kinases such as Shc or Src and thereafter interacts and activates the catalytic subunit p110 (39). The activated PI3-K subsequently activates downstream effector kinases such as Akt (PKB), a serine/threonine kinase. Akt has been shown to activate either directly or indirectly proteins responsible for prosurvival and anti-apoptosis in cancer cells. Several reports indicate that the PI3-K/Akt pathway interacts with the ERα pathway, resulting in bidirectional cross talk. It has been shown that PI3-K/Akt can mediate E2-induced transcription of cyclin D1 and entry of MCF-7 breast cancer cells into S-phase (40). In addition, the PI3-K/Akt pathway can induce ERα phosphorylation on Ser-167 to promote ERα-mediated transcription and thus protect cells against tamoxifen-induced apoptosis (41). The accumulating data suggest that ERα-positive breast tumors with alterations of PI3-K and Akt signaling, whether dependent or independent of EGFR/HER2/neu, might be insensitive to antiestrogens and thus lead to resistance. Future studies are needed to confirm whether blocking PI3-K and/or Akt signaling in ERα-positive breast cancer will be beneficial in subverting resistance to antiestrogens.

New SERMS

New drug discovery for selective ER modulators (SERMs) is currently driven by the known side effects of tamoxifen, which include an increase in the incidence of endometrial cancer, development of resistance, and the recent report of the negative effects of hormone replacement therapy (HRT) on coronary heart disease (42). Several approaches are being pursued by altering the antiestrogenic side chain or improving the pharmacokinetics of existing molecules. Several novel antiestrogen compounds have been developed to replace tamoxifen for ERα-positive breast cancer without the agonist actions. These compounds have the potential to be more effective than tamoxifen, having the advantage of reducing the incidence of endometrial cancer and possibly acquired resistance during therapy. Two groups of antiestrogenic drugs exist today with the potential to replace tamoxifen: (1) The SERMs that include tamoxifen-like compounds (triphenylethylenes) such as toremifene, idoxifene, and GW 5638 and fixed-ring compounds (benzothiophenes) such as raloxifene, arzoxifene, EM 652, and CP 336,156, and (2) selective ER downregulators (SERDs and pure antiestrogen) such as ICI 182,780 (Fulvestrant\textsuperscript{H}). In phase II trials in patients with tamoxifen-resistant metastatic breast cancer, the new SERMs showed low response rates (0–15%) (43,44), suggesting cross-resistance to tamoxifen. No difference in outcome was noted between tamoxifen and toremifene in patients with metastatic or early-stage breast
cancer (45,46). In contrast, few clinical trials exist today evaluating the effectiveness of the fixed-ring compounds to tamoxifen. The Study of Tamoxifen and Raloxifene (STAR) trial compared the efficacy of raloxifene over tamoxifen as a chemopreventive for women at high risk for breast cancer. Both drugs reduced the risk of invasive breast cancer by approximately 50%. Tamoxifen also lowered the incidence of lobular carcinoma in situ (LCIS) and ductal carcinoma in situ (DCIS) by about 50 percent; however, raloxifene did not have an effect on the incidence of LCIS and DCIS (47). Of note, the incidence of uterine cancer was 36% lower, and the incidence of thromboembolic events was 29% lower in the raloxifene arm (47). Fulvestrant and other SERDs have been shown to have high affinity for ERs compared with tamoxifen with none of the agonist activities. In clinical trials, fulvestrant initially showed significant promise for the treatment of advanced breast cancer (48,49). However, in a phase III clinical trial versus tamoxifen for first-line therapy of advanced breast cancer, fulvestrant was not superior to tamoxifen in terms of time-to-treatment failure (50). Therefore, more work is needed to evaluate the efficacy of SERDs in comparison with SERMs for ERα-positive breast cancer.

AROMATASE INHIBITORS

An alternate strategy to endocrine therapy, which specifically inhibits binding of E2 to the ER, is to inhibit the production of E2 by blocking the cytochrome p450 aromatase enzyme, the rate-limiting enzyme that converts androgens (i.e., testosterone and androstenedione) to estrogens (i.e., E2 and estrone) in the adrenal gland, surrounding stroma, and adipose tissue of the breast tumor. The main drugs of this type are aromatase inhibitors, which include Type I (steroidal) or Type II (nonsteroidal) (Fig. 2). Steroidal inhibitors are competitive-substrate mimics of androstenedione. These include formestane and exemestane, which are irreversible inhibitors that bind with high affinity to the binding site of aromatase and are converted to a covalently bound intermediate. Nonsteroidal inhibitors include the first-generation aromatase inhibitor aminogluthethimide and the second-generation compounds anastrozole and letrozole. All nonsteroidal aromatase inhibitors act by binding reversibly to the enzyme and competitively inhibiting binding of the substrate androstenedione. The benefits of using aromatase inhibitors over tamoxifen are believed to be the complete deprivation of E2 and thus better efficacy for ERα-positive breast cancer (51). Recent clinical data have clearly demonstrated that anastrozole (52), letrozole (53), and exemestane (54) are more effective than tamoxifen as first-line treatments in patients with metastatic breast cancer. On the basis of clinical results (52,53), currently both anastrozole and letrozole are approved by the Food and Drug Administration for first-line treatment of postmenopausal, ERα-positive advanced breast cancer. The data from advanced breast cancer trials provided the rationale to perform large-scale clinical trials to determine whether there is an...
Mechanisms of Resistance

Clinically, patients that relapse after a previous response to tamoxifen usually have a clinical response to aromatase inhibitors (59,60). These results strongly indicate that the ERz continues to be expressed and is functional in breast tumors that are resistant to antiestrogens. However, although estrogen deprivation treatment might be more effective than tamoxifen in delaying resistance, eventually resistance to aromatase inhibitors will also develop. To date, it is unclear...
whether similar mechanisms of actions that have been identified for tamoxifen resistance are also involved in resistance to aromatase inhibitors. The exact mechanisms contributing to aromatase inhibitor resistance has yet to be fully elucidated. However, in vitro studies have identified mutations within the aromatase gene that confers resistance to aromatase inhibitors (61). These mutations have not yet been identified in human breast carcinomas (62). Other studies have demonstrated that estrogen deprivation supersensitizes the breast cancer cell to low levels of estrogen, thus creating a hypersensitive environment to overcome estrogen deprivation resulting in resistance (63–65). In addition, results suggest that there is increased cross talk between growth factor receptor signaling pathways and ERz. ERz has been shown to become activated and supersensitized by several different intracellular kinases, including MAPKs, insulin-like growth factors, and the PI3-K/Akt pathway (41,66–69). Therefore, the data suggest that ERz continues to be an integral part of the breast cancer cell signaling pathway even after resistance to aromatase inhibitors has developed.

Role of Progesterone Receptor and HER2/Neu

There is emerging evidence to suggest that ER-positive cancers that do not express the progesterone receptor (PR) and/or HER2/neu are somewhat intrinsically resistant to tamoxifen and perhaps hormonal therapy in general. Arpino et al. have demonstrated an increased relapse rate in patients with ER-positive, PR-negative cancers compared with ER-positive, PR-positive cancers, treated with tamoxifen (70). Patients treated with tamoxifen with ER-positive, HER2-positive metastatic breast cancers have a shorter time to treatment failure compared with ER-positive, HER2-negative cancers (71). In fact, Arpino et al. have demonstrated an increase in both HER1 and HER2 in ER-positive, PR-negative cancers compared to ER-positive, PR-positive cancers, suggesting an interplay between the ER and epidermal growth factor pathways (70).

Two very small trials demonstrated a significantly increased clinical response rate in patients with ER-positive, HER2-positive cancers treated with preoperative aromatase inhibitors compared to preoperative tamoxifen (72,73). This led to a widely accepted hypothesis that aromatase inhibitors were a better choice than tamoxifen in patients with ER-positive, HER2-positive cancers. However, an analysis of the BIG-1-98 trial demonstrates that letrozole improves outcome compared to tamoxifen in both ER-positive, HER2-positive cancers (HR 0.68) and in ER-positive, HER2-negative cancers (HR 0.72) (74). A recent subanalysis of the ATAC trial demonstrated a significantly improved outcome in ER-positive, HER2-negative cancers (HR 0.66) but not in ER-positive, HER2-positive cancers (HR 0.92), but this may have been due to the small number of patients in the HER-positive group (75). Are HER2-positive cancers somewhat resistant to not just tamoxifen but also to aromatase inhibitors? As outlined above, a recent trial randomized patients with HR-positive, HER2-positive metastatic breast cancers to anastrozole alone or to anastrozole plus trastuzumab
Although there was no significant difference in overall survival, possibly because patients randomized to anastrozole alone could receive trastuzumab at disease progression, the time to progression was doubled from 2.4 months in the anastrozole-alone arm to 4.8 months in the combined arm ($p = 0.0016$) (76). The clinical benefit rate in the combination arm was $42%$—significantly higher than in the anastrozole alone arm. A trial that evaluated single-agent trastuzumab as first-line therapy for patients with HER2-positive cancers demonstrated a clinical benefit rate of $48%$ (77). This suggests the intriguing possibility that HR-positive, HER2-positive cancers are driven by the HER2 pathway, which renders the cancers partly resistant to hormonal therapies.

An initial evaluation of the ATAC trial using case report forms revealed that TTR was longer for anastrozole in both ER-positive/PR-positive and ER-positive/PR-negative subgroups, but the benefit was more pronounced in the ER-positive/PR-negative subgroup [HR 0.84, 95% confidence interval (CI) 0.69–1.02 vs. 0.43, 95% CI 0.31–0.61] (78). Importantly, the ER and PR analyses were not performed centrally. More recently, a central analysis of about 2000 patients on the ATAC trial demonstrated similar improvements with the use of anastrozole compared to tamoxifen, regardless of PR status (HR anastrozole vs. tamoxifen 0.72 for ER-positive, PR-positive subgroups and 0.66 for ER-positive, PR-negative subgroups) (75). In the BIG-1-98 trial, similar benefits for letrozole compared to tamoxifen were seen in the ER-positive/PR-positive and ER-positive/PR-negative subgroups (74).

On the basis of this data, decisions regarding whether to start a patient on tamoxifen or an aromatase inhibitor should not be based on PR or HER2 status. Further molecular profiling may help in the future in making decisions regarding optimal hormonal therapies.

Role of Gene Expression Analysis

Recent studies have shown that gene expression profiling of breast tumors can be used to predict outcome. The initial gene expression signatures were developed using snap-frozen tissue, limiting their clinical application (79). An assay (Oncotype DX, Genomic Health, Inc., Redwood City, California, U.S.) has been developed and validated in women with early-stage, node-negative breast cancer using fixed, paraffin-embedded tissue. The 21-gene panel includes genes involved with proliferation, invasion, and hormone response and classifies a tumor as being low, intermediate, or high risk of recurrence based on the level of expression (80). Additional studies have shown that a woman’s recurrence score also predicts response to chemotherapy: women with low recurrence scores derive little benefit from adjuvant chemotherapy, whereas women with high recurrence scores benefit significantly from chemotherapy in addition to hormonal therapy. Most strikingly are the findings that breast cancers with a high recurrence score obtain minimal, if any, benefit from five years of tamoxifen (81). Given the fact that these cancers likely have higher expression of HER2...
and lower quantitative hormone receptors, this is in keeping with the data outlined above, suggesting that HER2-positive cancers have intrinsic resistance to tamoxifen. It is not known if women with intermediate recurrence scores benefit from chemotherapy in addition to hormonal therapy (82). The TAILORx (Trial Assigning Individualized Options for Treatment) trial is under way, randomizing early-stage breast cancer patients with intermediate recurrence scores to chemotherapy plus hormonal therapy versus hormonal therapy alone.

FUTURE RESEARCH AND THERAPEUTICS

It is clear that the central regulator of growth for ER\(\alpha\)-positive breast tumors is ER\(\alpha\). ER signaling is evidently not isolated from other cellular signaling pathways. Research studies demonstrate that there is bidirectional cross talk between EGFR/HER2/neu, MAPK, PI3-K/Akt, and ER\(\alpha\), which ultimately affects the cellular response of a cell to estrogens and mitogens. ER\(\alpha\) integrates numerous signals from hormones, growth factors, and intracellular kinases along with the array of coregulators to modulate cellular physiology and tumor pathology. The first generation antiestrogen/SERM, tamoxifen, has been effective at increasing overall survival of patients with ER\(\alpha\)-positive breast cancer. However, the eventual development of resistance to tamoxifen is common because of the multiple signaling networks affecting the function of ER\(\alpha\). As a result of this limitation, intense research over the past 25 years has revealed the need for alternate treatment strategies to tamoxifen such as newer and better SERMs with less agonist activity in the uterus while being full antagonists in the breast.

In addition to SERMs, other therapeutic approaches have emerged such as the use of aromatase inhibitors to prevent the synthesis of E\(_2\) and thus deprive breast cancer cells and other cells of E\(_2\). The strategy of estrogen deprivation may also have consequences such as resistance, bone loss, and/or dementia. In the ATAC and BIG 1-98 adjuvant trials, changes in bone mineral density (BMD) and increased incidence of bone fractures were reported. Longer follow-up in these studies will help determine if the decrease in BMD stabilizes over time or if it continues throughout the course of treatment.

Therapeutic opportunities also exist from research studies investigating the role of growth factor receptors EGFR and HER2/neu and their intracellular signaling communicators MAPK and PI3-K/Akt in the development of tamoxifen resistance. Trastuzumab, the HER2/neu-specific antibody, has been shown to be highly effective in treating breast tumors overexpressing HER2/neu. Moreover, research data indicate a role for HER2/neu in antiestrogen resistance, demonstrating a rationale for using trastuzumab to treat or prevent antiestrogen-resistant breast cancer. EGFR inhibitors and other tyrosine kinase inhibitors may also prove to be beneficial in preventing resistance when given in combination with tamoxifen or other SERMs or SERDs such as fulvestrant.

In addition to EGFR and HER2/neu as important modulators of tamoxifen action, dissecting the specific interrelationship between ER\(\alpha\) and its coregulators
Individualization of Endocrine Therapy in Breast Cancer

may open the door for novel therapeutics. An understanding of how ERα-mediated gene transcription is dysregulated in breast cancer and how this changes in resistance during endocrine therapy will hopefully lead to improved strategies for the use of current and future antiestrogen therapies.

REFERENCES


35. Kurokawa H, Lenferink AE, Simpson JF, et al. Inhibition of HER2/neu (erbB-2) and mitogen-activated protein kinases enhances tamoxifen action against...


INTRODUCTION

There have been substantial technical and analytical advances within the past decade that have facilitated high-throughput analysis of clinical specimens, a process that has been referred to as “molecular profiling.” This term, as applied here, may refer to evaluating various “markers,” including genomic, proteomic, and epigenomic expression patterns, or a combination of these patterns, in clinical specimens. Technological advances have led to the ability to measure thousands of markers with a variety of validated methods (1), and analytical models have been developed that facilitate analysis of the voluminous amount of data that are generated (2). The technology has also led to an ability to perform “discovery-based research,” in which large volumes of data are generated and analyzed without a specific hypothesis, in contrast to the traditional scientific paradigm of “hypothesis-based research,” in which a limited number of genes or proteins are based upon a specific hypothesis and rationale (3). Discovery-based research and hypothesis-based research are not mutually exclusive; however, profiling may also be used to test specific hypotheses that are based on sound, scientific rationale. A series of studies have been reported over the past few years that have utilized gene expression profiling to discover “molecular markers,” which may identify (i) distinct molecular subtypes of breast cancer
(i.e., genotypic-phenotypic correlation), (ii) molecular signatures–associated prognosis (i.e., prognostic factors), and (iii) molecular signatures that predict the benefit from specific therapies (i.e., predictive factors) (4). When a “molecular marker” is shown to perform more reliably than clinical features in predicting prognosis or the benefit from a specific intervention, then there is interest in further evaluating the utility of the marker in clinical practice (5).

**PROGRAM FOR THE ASSESSMENT OF CLINICAL CANCER TESTS**

In order to address the issue of integrating molecular diagnostic testing into clinical practice, the U.S. National Cancer Institute initiated the program for the assessment of clinical cancer tests (PACCT) (http://www.cancerdiagnosis.nci.nih.gov/assessment/index.html). The goals of PACCT are to ensure translation of new knowledge about cancer and new technologies to clinical practice and development of more informative laboratory tools to help maximize the impact of cancer treatments. This effort has led to the publication of the REMARK (REporting recommendations for tumour MARKer prognostic studies) guidelines (6) and to the TAILORx (Trial Assigning IndividuaLized Options for Treatment) trial developed by the North American Breast Cancer Intergroup (http://www.cancer.gov/clinicaltrials/digestpage/TAILORx). In reviewing the rationale for design of the TAILORx trial in breast cancer, the following questions were considered by the Breast Cancer Intergroup and are reviewed herein:

- What therapeutic intervention should the marker be used to select?
- What patient population should be evaluated?
- Which marker should be evaluated?
- What trial design would integrate current knowledge about the marker and address gaps in our knowledge about the current marker and future potential markers?

**WHICH TREATMENT INTERVENTION?**

Therapeutic options are used in an adjunctive manner for patients with early-stage breast cancer after surgery, irradiation, endocrine therapy, chemotherapy, and trastuzumab (Table 1). In contrast to irradiation, endocrine therapy, and trastuzumab, there are no clinical factors that may be used to predict the benefit from adjuvant chemotherapy. Chemotherapy is usually recommended on the basis of the risk of recurrence (i.e., prognostic factors) and age (7). In contrast to other therapies, the treatment effect of chemotherapy is modest, acute toxicities are frequent, and the cost is high (8). Therefore, evaluation of a marker that has the potential for predicting benefit from chemotherapy would be desirable.
WHICH PATIENT POPULATION?

In 2006, approximately 124,000 women in the United States were diagnosed with estrogen receptor (ER)–positive breast cancer associated with negative axillary lymph nodes, accounting for about 50% of all breast cancer diagnosed each year and nearly 9% of all new cases of cancer (9). Treatment recommendations have generally been based on prognostic factors, with chemotherapy recommended if the residual risk of recurrence exceeds approximately 5% to 10% despite adjuvant hormonal therapy (i.e., the tumor size exceeds 1 cm or smaller tumors are associated with unfavorable histologic features) (7). By these criteria, adjuvant chemotherapy could be recommended annually for up to 75,000 women younger than 70 years in the United States. Currently, it is estimated that approximately 25% of all women, or about 30,000 annually, with early-stage ER-positive breast cancer receive adjuvant chemotherapy. Its use has increased substantially over the past 15 years (10) is highly dependent on age, and is estimated to be given to 75% of those younger than 50 years, 30% between the ages of 50 and 69 years, and about 5% of those 70 years or older.

Approximately 85% of women with early-stage ER-positive breast cancer are therefore adequately treated with adjuvant hormonal therapy. Although adding chemotherapy reduces the risk of recurrence on average by about 30%, the absolute benefit for an individual patient is small, ranging from 1% to 5% (11). The vast majority of patients with ER-positive breast cancer are therefore overtreated with chemotherapy, since most would have been cured with hormonal therapy alone. Decision aids such as Adjuvant! Online (12), decision boards (13), and other tools, are often useful to assist patients and caregivers with information regarding absolute benefits that might be expected from chemotherapy (14). Although such decision aids may assist some patients in making a more informed decision regarding whether to accept adjuvant chemotherapy, when faced with a choice, many patients and their clinicians err on the side of overtreatment because of the imprecise nature of predicting the treatment benefit.

### Table 1  Standard Adjuvant Therapies for Operable Breast Cancer

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Selection</th>
<th>Percent eligible</th>
<th>Treatment effect (%)</th>
<th>Acute toxicity</th>
<th>Relative cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemotherapy</td>
<td>Recurrence risk &gt; 5–10%</td>
<td>Up to 90</td>
<td>25–35</td>
<td>High</td>
<td>+++</td>
</tr>
<tr>
<td>Endocrine therapy</td>
<td>ER- and/or PR-positive tumor</td>
<td>70</td>
<td>50</td>
<td>Low</td>
<td>++</td>
</tr>
<tr>
<td>Trastuzumab</td>
<td>HER2/neu-positive tumor</td>
<td>15</td>
<td>50</td>
<td>Low</td>
<td>++++</td>
</tr>
<tr>
<td>Radiation</td>
<td>Breast-sparing surgery</td>
<td>40–60</td>
<td>90</td>
<td>Low</td>
<td>+</td>
</tr>
</tbody>
</table>

*Abbreviations: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.*
(15). Therefore, developing a clinical trial that employs a molecular marker to define the chemotherapy benefit was judged by the Breast Cancer Intergroup to be the preferred setting. Recent evidence demonstrating that the incremental benefits associated with contemporary chemotherapy regimens had minimal impact in high-risk ER-positive disease reinforces this approach (16).

WHICH MARKER?

A number of molecular diagnostic tests that provide more accurate prognostic information that standard clinical features have been developed and externally validated and are summarized in Table 2, including the Amsterdam 70-gene profile (17,18), Rotterdam 76-gene profile (19,20), and a 21-gene profile (21,22). Recently, these assays have been shown to have comparable discriminatory value in early-stage breast cancer (23).

The 21-gene Oncotype DX™ (Genomic Health, Inc., Redwood City, California U.S.) includes 16 tumor genes and 5 reference genes, with the result expressed as a computed “recurrence score” (RS) (Fig. 1) (21). The genes utilized in the assay were identified in several training set trials (24), followed by a validation study that was performed in patients with ER-positive, node-negative breast cancer who received a five-year course of tamoxifen (TAM) in National Surgical Adjuvant Breast and Bowel Project (NSABP) trial B-14 (21). A higher RS was associated with an elevated risk of distant recurrence, whether evaluated as a categorical or a continuous variable (Fig. 2). This assay was selected for evaluation in the TAILORx trial for several scientific and practical reasons that are summarized in Table 3 and discussed herein. In addition to the prospective validation study illustrated by Figure 2, a population-based external validation study performed in patients treated in the Kaiser Permanente health care system provided further evidence for the robustness of the assay (22). Moreover, subsequent studies in patients treated with TAM with or without

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Externally Validated Multigene Breast Cancer Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refs.</td>
<td>Number of genes</td>
</tr>
<tr>
<td>27,28</td>
<td>534</td>
</tr>
<tr>
<td>17,18</td>
<td>70</td>
</tr>
<tr>
<td>21,22</td>
<td>21</td>
</tr>
<tr>
<td>19,20</td>
<td>76</td>
</tr>
</tbody>
</table>

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; Chemo, chemotherapy; TAM, tamoxifen.
adjuvant chemotherapy in NSABP trial B20 indicated that only patients with an elevated RS (RS > 30) derived benefit from adjuvant chemotherapy (25). Other analyses also indicated that an RS also predicts local recurrence, which may occur in up to 20% of patients with operable breast cancer and which may be reduced by adjuvant chemotherapy (26). Furthermore, an RS more accurately predicted recurrence than the Adjuvant! Online model, providing evidence that it offers information that is complementary to Adjuvant! (12). Finally, although Oncotype DX includes a reverse transcriptase polymerase chain reaction (RT-PCR) assay-based evaluation of ER expression, a tumor with high ER content by RT-PCR may have either a low or a high RS, indicating that it reflects a measure other than simply ER expression (25).
There were also several practical reasons for selecting Oncotype DX. First and foremost is the fact that it involves an RT-PCR-based technique that may be employed on as little as three 10-μm sections of routinely processed and stored formalin-fixed, paraffin-embedded tissue. In addition, the assay was approved by Clinical Laboratory Improvement Amendments (CLIA) in the United States and has received favorable local coverage decisions regarding reimbursement by Medicare, indicating acceptance of the value of the assay by third-party payers. Finally, selecting Oncotype DX for further evaluation by the Breast Cancer Intergroup represents a logical extension of the successful public-private collaboration between NSABP and Genomic Health.

**WHAT TRIAL DESIGN?**

In designing a trial integrating a molecular diagnostic test into clinical decision making, three factors must be considered. First, it is essential to incorporate features already known about the prognostic and predictive value of the test into the trial design. In this regard, it is clear that patients with a very low RS do very well with hormonal therapy alone, and patients with a very high RS derive great benefit from chemotherapy, thereby providing sufficient information to make definitive treatment recommendations for patients who fall into these groups. Second, the trial must be designed to address the clinical utility of the test in circumstances where its utility is uncertain or result uninformative. This scenario applies to patients with a tumor that has a mid-range RS, in which the risk of relapse may be sufficiently high to recommend chemotherapy, but in which chemotherapy may be not be beneficial. Third, it is important to design the trial in a manner that offers an opportunity to evaluate other clinical cancer tests as they evolve, obviating the need to repeat a large-scale clinical trial and providing an opportunity to evaluate the test within a very short time frame.
The factors outlined in the previous paragraph were taken into account while designing TAILORx. Only patients who meet established clinical criteria for recommending adjuvant chemotherapy, who are medically suitable candidates for chemotherapy, and who agree to have their treatment assigned or randomized on the basis of the molecular test are eligible. After informed consent and “preregistration,” a tumor specimen is forwarded to Genomic Health for the Oncotype DX assay, with the result reported to the site within 10 to 14 days. After the Oncotype DX RS is known, patients are then registered and have their treatment assigned or randomized on the basis of the RS (Fig. 3), as described below:

- RS is 10 or less: Hormonal therapy alone
- RS is 26 or higher: Chemotherapy plus hormonal therapy
- RS is 11 to 25: Patients randomized to receive chemotherapy plus hormonal therapy (the standard treatment arm) versus hormonal therapy alone (the experimental treatment arm)

Because the study includes only women who meet standard clinical criteria for adjuvant chemotherapy and are medically suitable candidates for therapy, chemotherapy plus hormonal therapy is considered the standard treatment arm for patients with a mid-range RS (i.e., 11–25). Hormonal therapy is considered the experimental treatment arm for the mid-range group. The primary objective

Figure 3  TAILORx schema. Abbreviation: RS, recurrence score.
of the trial is to determine whether hormonal therapy alone is not inferior to chemotherapy plus hormonal therapy. The trial utilizes a “noninferiority” design for the mid-range group and is adequately powered to detect a 3% or greater difference between the two treatment arms.

The chemotherapy regimen and hormonal agents used are prescribed at the discretion of the treating physician but must be consistent with one of several standard options described in the protocol document. Patients may also enroll in other Intergroup trials as long as the protocol treatment is consistent with the TAILORx-assigned treatment arm (i.e., hormonal therapy or chemotherapy plus hormonal therapy). In addition, patients who have already had the Oncotype DX assay performed are eligible to enroll if the RS is 11 to 25.

At the time of registration and treatment assignment, patients are asked to provide blood samples for banking of plasma and peripheral blood mononuclear cells. Tissue specimens are collected by the Eastern Cooperative Oncology Group (ECOG) Pathology Coordinating Office, with central testing and confirmation of ER and progesterone receptor (PR) expression, creation of tissue microarrays, and RNA extraction for future studies.

The trial is available through the Cancer Trials Support Unit (www.ctsu.org) and is being coordinated by ECOG. Other participants include NSABP and all members of the North American Breast Intergroup, including the Southwest Oncology Group (SWOG), Cancer and Acute Leukemia Group B (CALGB), North Central Cancer Treatment Group (NCCTG), National Cancer Institute of Canada (NCIC), and the American College of Surgeons Oncology Group (ACOSOG).

**RATIONALE FOR THE RS RANGES IN TAILORx**

The RS ranges used in TAILORx are different from those originally described for the low- (<18), intermediate- (18–30), and high- (>30) risk groups as originally reported for the assay. The range was adjusted in order to minimize the potential for undertreatment in both the high-risk group (also called “secondary study group 2”) and the randomized group (also called the “primary study group”). When the NSABP B20 data were analyzed using the RS ranges used in TAILORx, the treatment effect of chemotherapy was similar to the original analysis (Table 4). For the analyses described in Table 4, disease-free survival (DFS) is defined as time to first local, regional, or distant recurrence, second primary cancer, or death due to any cause, and distant recurrence-free survival (DRFS) is defined as time to first distant recurrence or death from breast cancer (contralateral breast cancers, other second primary cancers, and deaths due to other cancers are censored, and local and regional recurrences are ignored). DFS has been the primary endpoint evaluated in most breast cancer trials and is the primary endpoint for TAILORx, whereas DRFS was the primary endpoint in the NSABP trials evaluating RS. Although a trend favoring the addition of chemotherapy becomes evident at an RS of approximately 11 when the risk of relapse is analyzed in a linear fashion, the 95% confidence intervals completely
### Table 4  RS and Response to Chemotherapy in B20 Trial ($n = 651$) by RS

<table>
<thead>
<tr>
<th>RS</th>
<th>Number of patients</th>
<th>TAM 10-yr DRFS (%)</th>
<th>TAM + chemo 10-yr DRFS (%)</th>
<th>HR (95% CI) for recurrence by addition of chemo</th>
<th>$P$-value</th>
<th>TAM 10-yr DFS (%)</th>
<th>TAM + chemo 10-yr DFS (%)</th>
<th>HR (95% CI) for recurrence by addition of chemo</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;11</td>
<td>177 (27%)</td>
<td>98</td>
<td>95</td>
<td>1.788 (0.360, 8.868)</td>
<td>0.471</td>
<td>77</td>
<td>85</td>
<td>0.605 (0.317, 1.153)</td>
<td>0.124</td>
</tr>
<tr>
<td>11–25</td>
<td>279 (43%)</td>
<td>95</td>
<td>94</td>
<td>0.755 (0.313, 1.824)</td>
<td>0.531</td>
<td>81</td>
<td>76</td>
<td>1.106 (0.671, 1.823)</td>
<td>0.691</td>
</tr>
<tr>
<td>&gt;25</td>
<td>195 (30%)</td>
<td>63</td>
<td>88</td>
<td>0.285 (0.148, 0.551)</td>
<td>&lt;0.0001</td>
<td>53</td>
<td>75</td>
<td>0.446 (0.270, 0.738)</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

**Abbreviations**: RS, recurrence score; HR, hazard ratio; CI, confidence interval; DRFS, distant relapse-free survival; DFS, disease-free survival; TAM, tamoxifen; chemo, chemotherapy (which included cyclophosphamide, methotrexate, 5-fluorouracil or methotrexate/5-fluorouracil).
overlap in the 11 to 25 RS range (Fig. 4). Finally, a high RS has also been associated with an elevated risk of local relapse, which may also be reduced by giving adjuvant chemotherapy (26). An RS of 11 is associated with a risk of both local and distant relapse of about 10%, a threshold that has been typically used for recommending adjuvant chemotherapy.

CONCLUSION

TAILORx represents a major step forward into the era of personalized medicine for breast cancer. By integrating a molecular diagnostic test into clinical decision making, patients and clinicians will be able to make more informed decisions regarding the most appropriate treatment options for a subset of patients for whom the test results in a clear treatment path and refine the utility of the assay for those for whom the result may be uninformative. This trial will also serve as an important resource for evaluating new molecular signatures and other technologies, such as proteomics, epigenomics, and pharmacogenomics, as these technologies evolve.

REFERENCES

Taking Prognostic Signatures to the Clinic

Michail Ignatiadis and Christos Sotiriou

Department of Medical Oncology, Translational Research Unit, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium

INTRODUCTION

The advent of microarray technology and the sequencing of the human genome have enabled scientists to simultaneously interrogate the expression of thousands of genes. These gene expression profiling studies have provided (i) a molecular classification of breast cancer into clinically relevant subtypes, (ii) new tools to predict disease recurrence and response to different treatments, and (iii) novel insight into various oncogenic pathways and the process of metastatic progression.

Breast cancer prognosis has been extensively studied by gene expression profiling, resulting in different signatures with little overlap in their constituent genes. In a large meta-analysis of all the above studies, despite the disparity in their gene lists, molecular subtyping and different prognostic signatures were shown to carry similar prognostic information. In all these signatures, proliferation genes were the common driving force with regard to prognostication (1).

Gene expression signatures, despite their current limitations may further refine breast cancer prognosis and prediction of benefit from systemic therapy, beyond conventional clinicopathological prognostic and predictive factors (2). In this chapter, we will not refer to gene expression signatures that predict response to systemic therapy. We will focus only on genomic predictors related to prognosis and the way these predictors are expected to impact patient management in the future and accelerate the transition from the “empirical” to “tailored” oncology.
MOLECULAR CLASSIFICATION OF BREAST CANCER BY GENE EXPRESSION PROFILING

Clinicians have long recognized that the diagnosis of breast cancer includes tumor types with different natural histories and responses to various treatments. Nevertheless, traditional histopathological characteristics have been unable to capture the biological heterogeneity of breast tumors. The first studies employing microarray technology to address this issue were performed by Perou and Sorlie (3–5). They proposed a molecular classification of breast cancer with at least five subgroups on the basis of the expression patterns of 500 “intrinsic genes.” Three subgroups were characterized by low to absent expression of the estrogen receptor (ER) and ER-related genes, the basal-like, the HER2++, and the normal breast-like subgroup. The first subtype was characterized by high expression of genes found in myoepithelial cells in the outer basal layer of a normal breast duct, like keratins 5/6 and 17, and was therefore named basal-like. The HER2++ subgroup was characterized by high expression of several genes of the ERBB2 amplicon. The normal breast-like subgroup showed genes expressed by adipose and nonepithelial tissues and also revealed a strong expression of basal epithelial genes. The other two subtypes consisted of ER-positive tumors and expressed genes of the luminal epithelial cells that are found in the inner layer of a normal breast duct, and, as a result, they were termed luminal. The luminal A subgroup had the highest expression of the ERα and the GATA-binding protein 3 gene, and the second subgroup (luminal B and C) showed low to moderate expression of these genes. To assign each sample to one of the five molecular subtypes (basal-like, HER2+/ER—, luminal A, luminal B, normal breast-like), Hu et al. have created a single sample predictor by calculating the mean expression profiles (centroids) of a new “intrinsic set of 300 genes” for each subgroup (6). In a multivariate analysis, the basal-like, HER2+/ER—, luminal A, and luminal B subgroups provided important prognostic information beyond those provided by traditional predictors like tumor size and grade.

Despite using a different microarray platform, Sotiriou et al. confirmed the consistency and the prognostic relevance of the above molecular classification of breast cancer in a different patient population (7). Again, ER status was the most important discriminator of gene expression patterns. Interestingly, although tumor grade was also strongly reflected in the transcriptome profile of the primary tumor, this was not the case for tumor size and axillary nodal invasion. These results confirmed the important role of ER and tumor grade biology in defining breast cancer molecular subtypes. In conclusion, microarray studies with different technology platforms and in different patient populations treated with heterogeneous treatments have been remarkably consistent in reproducing a similar molecular classification of breast cancer.

However, it should be pointed out that the number of robust breast cancer molecular subtypes is not certain, and there is no standardized method for assigning a molecular subtype to a new case (8). Nevertheless, the breast cancer molecular classification has begun to alter the way clinical trials are designed. As an example,
new trials have been designed to test the efficacy of targeted agents only in the basal-like breast tumors.

GENE EXPRESSION SIGNATURES AS TOOLS FOR IMPROVING PROGNOSIS IN BREAST CANCER

Besides the above studies aiming at providing a molecular taxonomy of breast tumors, several independent groups, including our own, have conducted gene expression profiling studies with the objective of improving upon traditional prognostic markers used in the clinic for prediction of outcome. Two conceptually different, supervised approaches for prognostic marker discovery on a genome-wide scale have been applied so far, the “top-down” approach and the “bottom-up” or “hypothesis-driven” approach. The former derives from a prognostic model simply by looking for gene expression patterns associated with the clinical outcome without any a priori biological scenario, whereas the latter approach first identifies gene expression profiles linked with a specific biological phenotype and subsequently correlates these findings to survival.

Using the top-down approach, researchers from the Netherlands Cancer Institute (NKI) first identified a 70-gene prognostic signature in a series of 78 systemically untreated node-negative breast cancer patients. This signature (MammaPrint® Agendia, Amsterdam, the Netherlands) was then validated on a larger set of 295 young treated and untreated patients, with either node-negative or -positive breast tumors, from the same institution (9,10). The Rotterdam group later identified, using a training set of 115 breast cancer patients, a 76-gene signature to predict the development of distant metastases in untreated node-negative breast cancer patients (11). Interestingly, this group used the ER status to divide the training set into two subgroups, and the genes selected from each subgroup (60 and 16 genes for the ER-positive and ER-negative subgroups, respectively) were combined to form a single signature. The 76-gene signature (VDX2 array, Veridex LLC, Warren, New Jersey, U.S.) was recently validated in 180 lymph node–negative, untreated patients from multiple institutions (12). The two groups have used different technology platforms (the NKI group used the Agilent platform, Santa Clara, California, U.S. and the Rotterdam group used the Affymetrix platform, Santa Clara, California, U.S.), and although there was only a three-gene overlap between them, both signatures outperformed the National Institutes of Health (NIH) (13) and St. Gallen criteria (14) in stratifying patients according to risk of relapse. Indeed, while the prognostic signatures correctly identified those patients who will relapse, they could also identify many low-risk patients who could be spared from unnecessary chemotherapy. However, the identification of high-risk patients could still be improved, since half of the patients assigned to the poor-signature group, in fact, did not relapse.

A major challenge for gene expression profiling studies, especially those with clear clinical implications, is independent validation. Consequently, TRANSBIG, the translational research network of the Breast International Group
(BIG), conducted an independent validation study of these two prognostic signatures in a series of 302 untreated patients from five different centers and across different statistical facilities (15,16). The two signatures were robust, and their performance was reproducible and similar in this retrospective series. Interestingly, both signatures were strong predictors of the development of metastasis within five years of diagnosis (early metastasis), but their performance decreased for the detection of late metastasis. This could be easily explained, since the training set in both studies consisted of patients who have all relapsed within five years. However, an intriguing question arises about the potential differences in the biology and the oncogenic pathways that drive early and late metastasis.

Another group explored the correlation between the expression of 250 genes (selected from publicly available microarray data linked with breast cancer biology) and the clinical outcome in 447 patients from three studies including the tamoxifen-only group of the NASBP trial B-20 to build the Oncotype DX<sup>1</sup> (Genomic Health Inc, Redwood City, California, U.S.) Recurrence Score<sup>1</sup> (RS) (17). This score (low, intermediate, and high) was based on the expression of a final selected panel of 16 cancer-related genes and five reference genes, as assessed by reverse transcriptase-polymerase chain reaction (RT-PCR) in paraffin-embedded tumor tissue. Then, the Oncotype DX RS was validated for its ability to quantify the likelihood of distance recurrence in the tamoxifen treated, ER-positive, node-negative group of the NSABP trial B-14 (17). Using the RS, around 50% of the patients in the above study were characterized as low risk (low RS), with a 6.8% probability of distant recurrence at 10 years.

Using the second approach, the hypothesis-driven approach, Chang et al. tried to identify the molecular similarities between cancer invasion/metastasis and wound healing (18). They developed a gene signature to characterize the response of fibroblasts to serum. When they applied these genes to primary gene expression profiling data from multiple data sets, they observed that the presence of a wound-healing phenotype in the primary tumor was related with worse clinical outcome in various cancers, including breast cancer. Similar to the 70- and 76-gene signatures, the wound-response signature provided better risk stratification than the NIH and St. Gallen criteria. Furthermore, it was able to refine prognosis in the 295 patient data sets of van de Vijver et al. (10) beyond existing clinicopathological risk factors and the 70-gene signature (19). In conclusion, Chang et al., using a hypothesis-driven approach, apart from identifying a gene expression prognostic signature, have offered novel insight into the molecular pathogenesis of metastasis by linking this process with wound healing.

Another example of deriving a prognostic gene expression signature using a hypothesis-driven approach is the study by our group (20) that focused on histological grade, a well-established pathological parameter rooted in the cell biology of breast cancer. Applying a supervised analysis, we developed a Gene expression Grade Index (GGI) score based on 97 genes. These genes were mainly involved in cell cycle regulation and proliferation and were consistently
Taking Prognostic Signatures to the Clinic

differentially expressed between low- and high-grade breast carcinomas. Interestingly, the GGI, which essentially quantifies the degree of similarity between the tumor expression pattern of these 97 genes and tumor grade, was able to reclassify patients with histological grade 2 tumors into two groups with distinct clinical outcomes similar to those of histological grades 1 and 3, respectively (20). This observation challenges the existence and clinical relevance of an intermediate grade classification. When we explored the implications of the joint distribution of ER status and GGI in predicting clinical outcome, we found that almost all ER-negative tumors with poor clinical outcome were associated with a high GGI score (high grade), whereas ER-positive tumors were associated with a heterogeneous mixture of gene expression grade values. Recently, our group has reported that GGI can be used to define two clinically relevant molecular subtypes in ER-positive breast cancer (21). These two ER-positive molecular subgroups (high and low genomic grade) were highly comparable to the luminal A and B classifications and were associated with distinct clinical outcome in systemically untreated or tamoxifen-only-treated ER-positive breast cancer populations involving more than 650 patients (21). GGI appeared to be the strongest predictor of clinical outcome, highlighting the prognostic importance of proliferation genes in ER-positive subgroups, as have been already reported (17). Indeed, proliferation was one of the five components associated with the highest weight in the Oncotype DX RS.

Four additional studies have used the hypothesis-driven approach to build gene predictors of clinical outcome. Oh et al. used estrogen-induced genes, identified by treating the MCF-7 cell line with 17β-estradiol to build a predictor of clinical outcome on a training set of 65 patients with ER-positive breast tumors (22). A total of 822 genes, mainly proliferation and antiapoptosis genes could separate these tumors into two subgroups with different clinical outcomes. This gene list was then mapped in three published datasets (4,5,19,23) and an average expression index or “centroid” was created that could classify the ER-positive tumors into two clinically relevant subgroups. In another study, Miller et al. generated a 32-gene expression signature that distinguishes p53 mutant and wild type tumors. The aim of this research was to provide a more accurate assessment of the functional status of the p53 pathway than that provided by mutation analysis using sequencing (24). It should be noted that none of the signature genes were known transcriptional targets of p53, nor have they been related to the p53 pathway. Instead, they were mainly proliferation-related genes, various transcription factors, and ion transport genes. Once more, the proliferation genes were a prominent part of this signature. The signature outperformed p53 mutation status, as identified by sequencing in prognosis and prediction of response to hormonal therapy and chemotherapy. By combining the hypothesis-driven approach with comparative cross-species functional genomics, Glinsky et al. identified an 11-gene “death-from-cancer” signature (25). This signature can identify a lethal subset of human cancers of various origins (including breast cancer) with a high propensity to metastasize. These tumors display the
“phenotype” of an activated BMI-1 oncogenic pathway that is essential for the self-renewal of normal stem cells. Again, proliferation- and cell cycle–regulating genes comprise the most essential part of this signature. Along the same lines, Liu et al. isolated and compared at the transcriptional-level tumorigenic CD44+/CD24−/low breast cancer cells derived from a small set of six breast cancer patients with normal breast epithilium (26). They developed an “invasiveness” gene signature (IGS) that was associated with the risk of death and metastasis not only in breast cancer but also in other tumor types. Again, there was a small gene overlap between IGS and other prognostic signatures, and, not surprisingly, its prognostic power was significant only in ER-positive and intermediate-grade tumors, a feature that had already been observed for other prognostic profiles.

CAN WE INTEGRATE THE DIFFERENT GENE EXPRESSION SIGNATURES?

The above gene expression signatures developed to improve breast cancer prognostication have little overlap in their constituent genes. A representative list of these signatures is depicted in Table 1. The inevitable question is whether the “intrinsic subtypes” developed by unsupervised cluster analysis and the other prognostic signatures have also little overlap in the prognostic information they convey. Fan et al. recently compared the prognostic ability of five gene expression–based models—intrinsic subtypes (6), 70-gene profile (9), wound response (19), RS (17), and two-gene ratio (23)—in a single dataset of 295 patients and found that four of five predictors had significant agreement in outcome predictions for individual patients (27). However, this study was limited to only one dataset of 295 patients, examined only five gene expression signatures, and did not investigate the biology that may drive prognosis among the five signatures. To address this issue, a large meta-analysis of publicly available gene expression and clinical data from 2833 patients (4–7,9,12,15,20,23,24,28–39) with breast cancer was performed by the Swiss Institute of Bioinformatics in collaboration with our group (Table 1). In this meta-analysis, the concept of “co-expression modules” (comprehensive list of genes with highly correlated expression) was used to describe important biological processes in breast cancer, like proliferation, ER and HER2 signaling, tumor invasion, and immune response. In this study, we sought to depict the connection between these modules, the previously reported molecular classification, several gene prognostic classifiers, and the most established clinicopathological variables. A number of interesting conclusions were drawn from this collaborative effort. First, the disparity of gene lists of the various gene signatures may be attributed to (i) heterogeneity in patient populations studied, (ii) the use of different microarray platforms with different probe sets and different methods for data normalization, and (iii) sampling variation due to small sample size relative to the number of genes examined. Second, using these co-expression modules, instead of the five molecular portraits reported with the intrinsic genes, breast tumors were now grouped into three main subgroups namely ER−/HER−, HER2+, and ER+/HER2−. Third, the ER−/HER−,
<table>
<thead>
<tr>
<th>Gene expression signatures</th>
<th>Biological scenario</th>
<th>Microarray platform</th>
<th>Number of genes in the signature</th>
<th>Independent validation</th>
<th>Prospective clinical validation</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>70-gene signature</td>
<td>Clinical outcome</td>
<td>Agilent (oligonucleotides)</td>
<td>70</td>
<td>yes</td>
<td>Yes (MINDACT Trial)</td>
<td>(9,10,15)</td>
</tr>
<tr>
<td>76-gene signature</td>
<td>Clinical outcome</td>
<td>Affymetrix (oligonucleotides)</td>
<td>76</td>
<td>yes</td>
<td>No</td>
<td>(11,12,16)</td>
</tr>
<tr>
<td>Recurrence Score</td>
<td>Clinical outcome</td>
<td>RT-PCR</td>
<td>21</td>
<td>yes</td>
<td>Yes (TAILORX trial)</td>
<td>(17,37)</td>
</tr>
<tr>
<td>Wound-response signature</td>
<td>Wound healing and tumor progression</td>
<td>CDNA (custom made)</td>
<td>512</td>
<td>yes</td>
<td>No</td>
<td>(18,19)</td>
</tr>
<tr>
<td>Genomic grade</td>
<td>Histological grade and tumor progression</td>
<td>Affymetrix (oligonucleotides)</td>
<td>97</td>
<td>yes</td>
<td>No</td>
<td>(20)</td>
</tr>
<tr>
<td>P-53 signature</td>
<td>Functional status of p-53</td>
<td>Affymetrix (oligonucleotides)</td>
<td>32</td>
<td>yes</td>
<td>No</td>
<td>(24)</td>
</tr>
<tr>
<td>Death-from-cancer signature</td>
<td>Bmi-1 oncogenic pathway self-renewal</td>
<td>Affymetrix (oligonucleotides)</td>
<td>11</td>
<td>Yes</td>
<td>No</td>
<td>(25)</td>
</tr>
<tr>
<td>Invasiveness gene signature</td>
<td>Tumorigenic cancer cells CD44+/CD24−, low</td>
<td>Affymetrix (oligonucleotides)</td>
<td>186</td>
<td>Yes</td>
<td>No</td>
<td>(26)</td>
</tr>
</tbody>
</table>
HER2+ tumors were characterized by high proliferation, whereas the ER+/HER2− was divided in two subgroups, the ER+/low proliferation and the ER+/high proliferation tumors resembling to luminal A and B subtypes, respectively. Fourth, 10 prognostic signatures [RS (17), 70-gene signature (9), 76-gene signature (11), wound response (18), p53 signature (24), GGI (20), NCH70 (33), CON52 (38), CCYC (39), ZCOX], despite the disparity in their gene lists, showed similar prognostic performance (Fig. 1). It is interesting to note that they all identified as “low risk” the ER+ subtype associated with low proliferation. Fifth, proliferation was the common driving force responsible for the performance of the various prognostic signatures (Fig. 1). Indeed, to further investigate the role of proliferation genes in relation to the different prognostic signatures, we divided each signature into two “partial signatures”—one with only proliferation genes and the other with the complementary nonproliferation genes. Interestingly, when only proliferation genes were used, the total performance was not degraded. In contrast, the nonproliferation partial signatures typically show degraded performance. Sixth, combining the signatures did not improve the performance as expected from the high concordance in their classifications. Finally, in the multivariate analysis, including the genomic predictor derived from co-expression modules, the standard clinicopathological variables, namely tumor size and nodal status, retained their prognostic value.

PROSPECTIVE VALIDATION OF TWO GENOMIC PREDICTORS

The initial promising results that breast cancer prognostication can be improved using gene expression signatures need to be validated by well-designed prospective clinical trials. Two gene expression predictors, the 70-gene expression
signature, (MammaPrint) discussed earlier and the Oncotype DX RS have reached the final step of prospective clinical testing. The BIG and the EORTC have jointly launched the MINDACT (Microarray for Node-Negative Disease may Avoid Chemotherapy) trial that directly compares two approaches for deciding whether to offer adjuvant chemotherapy in node-negative breast cancer patients, i.e., either using the prognostic information provided by standard clinicopathological criteria (control arm) or by the 70-gene expression signature (MammaPrint) (experimental arm). Six thousand women will participate, and the primary endpoint is to prove that five-year disease-free survival in the experimental arm will not be inferior to that of the control arm; at the same time, it is estimated that 10% to 15% fewer women will be treated with chemotherapy in the experimental arm.

A similarly ambitious initiative is taking place in the United States. The Oncotype DX RS is being evaluated in the TAILORx [Trial Assigning IndividuaLized Options for Treatment (Rx)] trial, which is part of the National Cancer Institute (NCI) Program for the Assessment of Clinical Cancer Tests (PACCT). This trial randomizes patients with intermediate RS to receive either hormonal therapy alone or hormonal therapy plus chemotherapy. In it, patients with low RS will be treated with hormonal therapy alone and those with high RS will receive chemotherapy plus hormonal therapy.

These two trials will provide level I evidence that these gene expression predictors provide prognostic information beyond standard clinicopathological factors and are ready for use in clinical practice. Both trials will provide an excellent opportunity to address issues related to tissue handling and shipping, reproducibility, quality controls, and standardization of the microarray experiments. Furthermore, they will provide a valuable database for future translational research.

FUTURE DIRECTION–CONCLUSIONS

The real challenge, once the above trials prospectively validate the two genomic predictors, will be to integrate these predictors into a comprehensive clinical decision-making algorithm that will lead toward personalized medicine. Parameters that need to be included in this algorithm include the following:

Traditional Clinicopathological Parameters

These include patient comorbidities, age, and primary tumor characteristics like tumor size, tumor grade, nodal status, ER status, HER2/neu status, and lymphovascular invasion. Many of these parameters are incorporated in various prognostic tools like the Nottingham Prognostic Index (40), Adjuvant! Online (41), and the NIH (13) and St. Gallen (42) Guidelines.
Primary Tumor Gene Expression Signatures

From the above discussion, it is apparent that two genomic predictors are in the late phase of clinical validation. However, apart from the question of who can be spared unnecessary adjuvant chemotherapy (prognosis), a more important question is which therapy works better in an individual patient (prediction). Genomic predictors of response to multidrug regimens and newer targeted agents have also been reported (43–49), but their description is beyond the scope of this chapter. Other important technology platforms are being developed to analyze epigenetic changes in DNA (50), the role of microRNAs (51) that will provide us with a more comprehensive picture of the biology of the tumor and further refine prognosis and prediction.

Micrometastatic Cells

The presence of cytokeratin positive tumor cells in the bone marrow of early breast cancer patients has been correlated with worse outcome in a meta-analysis involving more than 4500 patients (52). We have reported that the detection with a real-time RT-PCR assay of Cytokeratin-19 (CK19) mRNA-positive circulating tumor cells (CTCs) in patients with early breast cancer before (53,54) and after (55,56) the initiation of adjuvant systemic treatment is associated with poor outcome. However, the additional prognostic information of micrometastatic cells to primary tumor gene expression profiling data has not yet been examined. Moreover, there are pilot studies where HER2-positive CTCs and/or disseminated tumor cells (DTCs) were successfully eliminated with trastuzumab (57), suggesting that CTCs and/or DTCs may be tested also as tools to predict response to a particular drug.

Molecular Imaging

PET (positron emission tomography) scan using novel PET tracers that target critical cellular processes like angiogenesis may provide important staging and functional information and is under investigation in breast cancer (58).

Pharmacogenetics (Host)

The importance of pharmacogenetics in predicting clinical outcome in breast cancer has recently been illustrated in studies focused on the polymorphism of enzymes that metabolize tamoxifen (59,60). The study of the host is an important piece of the clinical decision-making algorithm.

Proteomics

Proteomics provides the hope of discovering novel biological markers that can be used for early detection, disease diagnosis, prognostication, and prediction of response to therapy in early breast cancer (61).
Adjuvant Treatment

It is well documented that adjuvant treatment affects clinical outcome of breast cancer patients (62). From that perspective, the positive adjuvant trials with the aromatase inhibitors (63) and trastuzumab (64,65), the concept of dose-dense adjuvant chemotherapy (66), and the emerging data about the use of taxanes (67,68) and anthracyclines (69) in the adjuvant setting are rendering both the treatment decision-making process and prediction of outcome increasingly more complex.

Several efforts have been made to combine at least some of the above parameters to predict the clinical outcome of breast cancer patients. For instance, Pittman et al. have used a modeling approach based on classification trees to combine clinical and genomic data (70). They nicely showed that by combining multiple metagenes (defined as summary measures of the expression of subsets of potentially related genes) with nodal status, they could create integrative clinicogenomic models. These models provided improved prediction accuracy compared with either clinical or genomic data alone. Another example of such an effort is the Adjuvant! Online for breast cancer, Genomic Version 7.0, which combines traditional clinical data, the RS, and information about adjuvant treatment to provide an estimate of clinical outcome in patients with early breast cancer (71).

However, the accumulation of highly complex and multilevel information will ultimately impose specific data integration requirements. So far these different types of data are not only maintained in a variety of different data sources but are also managed by different complex data management and analytical systems. Additionally, they involve many institutions and numerous complex workflows. To address all these critical issues, two initiatives have been launched that intend to integrate all this information into one comprehensive biomedical platform—the cancer Biomedical Informatics Grid, “caBIG” (https://cabig.nci.nih.gov/) and the Advancing Clinico Genomic Trials, “ACGT” (http://www.eu-acgt.org/) funded by the NCI and European Union, respectively.

In the future, medical oncologists should use a clinically validated treatment decision-making algorithm that will take into account all this highly complex and rapidly evolving, multilevel information. In this way, we will move from the “empirical” to the new exciting era of “individual-tailored” oncology.

REFERENCES

Taking Prognostic Signatures to the Clinic

INTRODUCTION

All eyes are on tissue. The biggest challenge facing biomedical discoveries today is the issue with tissue. The era of personalized, individually tailored medicine has begun; therefore, cells, tissue and its intra- and interenvironment, are the focus of pharmacogenetics. The intracellular kinetic and intercellular dynamics are key factors in the understanding of cell behavior before, during, and after therapy.

Undoubtedly the solid type of tissue carries the bulk of these issues. Other fluid forms of tissue, however, do have their own share of these issues, of course. Tissue will be defined here as biological material comprising a network of characteristically diverse cells. In essence, blood, urine, surgical resections, biopsies, marrow, lavage, hair, nail, and other specimen taken from a person will be considered as tissue. Turn to any biomedical community member and most, if not all, will have some sort of opinion about tissue. Between the ultimate giver (patients) and ultimate user (laboratorians) lie the challenges. To the patient, care team, and several entities, the tissue is the issue, and to the laboratories, technologists, and the interpreters of results, the issue is really the tissue.

In the past two years, several working groups were established not only to understand current issues as well as future anticipated issues but also to address them. These working groups are usually comprise various clinical and financial institutions and experts from multiple disciplines, such as medical scientists, basic scientists, surgeons, laboratories, pathologists, funding agencies, agent
manufacturers, patient advocates, statisticians, bioinformatics, research associates, research administrators, medical staff, and clinical administrators. These groups guide the interdisciplinary patient care team by addressing the procurement, collection, preservation, acquisition, processing, transportation, and storage of biological materials of high quality and usability that are adequate for analysis and interpretation. Good news is that specimens with clinical data are archived as far back as 40 years ago, the bad news is that they were collected non-uniformly; today’s good news is that the collection processing and associated data are undergoing harmonization (uniformity). Contingency plans are also being developed for future anticipated issues. The key areas of the tissue issue are obviously procurement, collection, acquisition, processing, preservation, storage, distribution, and bioinformatics that are successfully conducted within stringent policies, procedures, rules, and regulations.

PROCUREMENT

A strategy for collecting any biological specimens from patients is needed to ensure that the behavior of cells or tissue after it leaves the body of the patient is consistent with what it was while in the patient’s body. Regardless of the strategy, clear and simple instructions should be written and made available to the procurement team comprising a preoperation nurse, phlebotomist, surgery assistant, physician/surgeon, physician assistant, pathology assistant, research coordinator, patient care nurse, internal specimen transporters, and medical/research technologist. These instructions are typically included in the local manual of operational procedures and/or clinical protocols and should provide details of the type of collection devices such as vacutainers, syringes, punch size, preservatives, buffers, reagents and the containers needed. The manual must also include procedures for preparing reagents or buffers such as formalin, RNAlater, and the optimal cutting temperature (OCT) compound. It must contain information about the precollection requirements with regard to time, temperature, and pressure for certain types of specimens (Table 1).

COLLECTION

There are several choices of collection devices, to the extent that they have created a great deal of confusion in the medical field. It is not surprising to see several new clinical trial protocols now restricting collection to the use of preassembled collection devices, with instructions to assist the procurement plan and ensure uniform collection, particularly in a multi-institutional study. While it is possible to obtain serum from redtop, royal blue top, serum separator top, and plain syringe, it is true that tissue can be fixed in paraformaldehyde, formalin, acid alcohol, zinc formalin, and many other reagents. Each of these fixatives presents its own advantages and disadvantages toward downstream applications such as immunohistochemistry (IHC), in situ hybridization (ISH),
comparative genomic hybridization (CGH), polymerase chain reaction (PCR), blotting (Western, Northern, Southern), enzyme-linked immunosorbent assay (ELISA), radio immunoassay (RIA), enzyme immunoassay (EIA), and chemiluminescence assay (CLA). It is also true that plasma can be obtained from all anticoagulated vacutainers as well as heparin-coated syringes. Protein and nucleic acids can be obtained from biological samples regardless of collection or preservation buffer. However, experience has shown a great deal of variation in the yield and quality of sample products, as well as the analytical results obtained from the samples. It is therefore important to select one device and one alternate backup for the collection, preservation, and fixation that will provide the best representation of the cellular content and its content behavior as it was just before leaving the patient’s body. It is important to select a device and preservative that are best for yield and quality, as well as most appropriate for the next step in processing or analysis. Table 2 lists the collection devices, preservatives, and fixatives that have been selected by large research groups because of their overall performance success.

**Fresh tissue.** Nonfrozen and nonfixed tissue should be collected directly into a labeled sterile screw-cap plastic container, with one of the following buffers:

1. Phosphate buffered saline (PBS) for tissue culture and cell viability studies.
2. Neutral buffered formalin (10%) for gradual fixation.

<table>
<thead>
<tr>
<th>Roles</th>
<th>Responsibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician/surgeon including but not limited to pathologist and intervention radiologist</td>
<td>Inform patient, order test, perform surgery/biopsy/imaging/morphology/assay procedures.</td>
</tr>
<tr>
<td>Nurse</td>
<td>Inform patient, determine timing of procedure/draw. Draw blood via port.</td>
</tr>
<tr>
<td>Phlebotomist, nurse, medical technologist</td>
<td>Draw blood into devices. Perform aspiration.</td>
</tr>
<tr>
<td>Physician/surgery assistant</td>
<td>Guide physician and expedite preservation and delivery.</td>
</tr>
<tr>
<td>Transporter</td>
<td>Ensure timely delivery from point of collection to next critical step/location while preserving samples.</td>
</tr>
<tr>
<td>Research coordinators, project manager, protocol specialist, data manager, nurse coordinator, clinical research associates</td>
<td>Ensure all team members are familiar with their roles and have adequate information to assist in performing task. Liaison among team members as well as facilitate collaboration.</td>
</tr>
<tr>
<td>Medical/research technologist</td>
<td>Ensure availability of appropriate supplies for collection and preservation. Effectively and efficiently process specimens and analyze samples.</td>
</tr>
</tbody>
</table>
### Table 2  Common Blood Collection Tubes

<table>
<thead>
<tr>
<th>Closure/stopper color</th>
<th>Anticoagulant or additive</th>
<th>Number of inversions to mix</th>
<th>Examples for use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red/red in glass tube</td>
<td>None</td>
<td>None</td>
<td>Serum for chemistry, proteomics, immunology, and banking</td>
</tr>
<tr>
<td>Red/red in plastic tube</td>
<td>Clot activator</td>
<td>5–10</td>
<td>Serum for chemistry, proteomics, immunology, and banking</td>
</tr>
<tr>
<td>Gold/red marbled SST</td>
<td>Clot activator and separation gel</td>
<td>5–10</td>
<td>Serum for chemistry, proteomics, immunology, and banking. Remote collection.</td>
</tr>
<tr>
<td>Lavender/lavender (purple)</td>
<td>Sodium EDTA (dried), or potassium EDTA (liquid)</td>
<td>5–10</td>
<td>Whole blood for hematology, certain whole blood chemistries, e.g., hemoglobin A1c or blood typing. Proteomics, floating DNA, genomic DNA, cells.</td>
</tr>
<tr>
<td>Light blue/light blue</td>
<td>Liquid 3.2% sodium citrate, or 3.8% sodium citrate</td>
<td>5–10</td>
<td>Plasma for coagulation, Platelet poor plasma. Growth factors, fragile proteins, viable cells.</td>
</tr>
<tr>
<td>Cell preparation tube CPT</td>
<td>Liquid 3.2% sodium citrate, or 3.8% sodium citrate, Ficol and gel separator</td>
<td>5–10</td>
<td>Plasma for coagulation, Platelet poor plasma. Growth factors, fragile proteins, viable cells. Remote collection of viable mononuclear cells.</td>
</tr>
<tr>
<td>Paxgene RNA/DNA</td>
<td>Preanalytic/Qiagen nucleic acid stabilizer</td>
<td>5–10</td>
<td>Genomic expressions.</td>
</tr>
<tr>
<td>P100</td>
<td>Protein stabilizer</td>
<td>5–10</td>
<td>Small size protein measurement. Proteomic studies.</td>
</tr>
<tr>
<td>Tan</td>
<td>Sodium heparin (glass), or potassium EDTA (plastic)</td>
<td>5–10</td>
<td>Whole blood for lead, heavy metals, toxicology studies.</td>
</tr>
<tr>
<td>Closure/stopper color</td>
<td>Anticoagulant or additive</td>
<td>Number of inversions to mix</td>
<td>Examples for use</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------</td>
<td>-----------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Green/green</td>
<td>Lithium heparin, or</td>
<td>5–10</td>
<td>Plasma or whole blood for chemistry, PK, and PD.</td>
</tr>
<tr>
<td></td>
<td>Sodium heparin</td>
<td></td>
<td>Plasma for chemistry</td>
</tr>
<tr>
<td>Light green/green</td>
<td>Lithium heparin and separation gel</td>
<td>5–10</td>
<td>Plasma for chemistry</td>
</tr>
<tr>
<td>marbled</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gray/gray</td>
<td>Potassium oxalate/sodium fluoride, or sodium fluoride, or lithium iodoacetate, or lithium iodoacetate/lithium heparin</td>
<td>5–10</td>
<td>Serum or plasma for glucose or alcohol requiring inhibition of glycolysis, C-peptides.</td>
</tr>
<tr>
<td>Dark royal blue</td>
<td>Sodium heparin, EDTA, or plain</td>
<td>5–10</td>
<td>Plasma or serum for trace elements, toxicology, and nutrients.</td>
</tr>
<tr>
<td>Pale yellow/bright yellow</td>
<td>Sodium polyanetholesulfonate</td>
<td>5–10</td>
<td>Blood cultures, microbiology, and infectious diseases markers.</td>
</tr>
<tr>
<td>Yellow</td>
<td>Acid citrate dextrose (ACD)</td>
<td>5–10</td>
<td>Blood bank, viable cells alternate to CPT and CellSave Serum for chemistry, ELISA</td>
</tr>
<tr>
<td>Orange/yellow and</td>
<td>Clot activator (thrombin)</td>
<td>5–10</td>
<td></td>
</tr>
<tr>
<td>marbled</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pink/pink</td>
<td>Potassium EDTA</td>
<td>5–10</td>
<td>Molecular/viral load testing, blood bank, hematology.</td>
</tr>
</tbody>
</table>

<sup>a</sup>www.bd.com  
*Abbreviations*: SST, serum separator top; PK, pharmacokinetic; PD, pharmacodynamic.
3. RPMI (Roswell Park Memorial Institute) media for tissue culture, cell viability studies, protein extraction, and multiphoton studies.
4. RNA later for RNA extraction.
5. OCT bottom and top prior to freezing.

PROCESSING

Generally, processing methods that will ensure optimal quality and yield should be selected for the processing of specimens and samples into next product aliquot.

**Blood.** It is important to perform all centrifugation in refrigerated and temperature-controlled chambers. This is to avoid unnecessary heat incubation of samples while spinning. In some instances, it will be necessary to spin the sample at a slow speed followed by high speed to eliminate platelets. Selection of complex processes to obtain delicate and fragile by-products is important. Aliquoting into the smallest useable quantity to avoid freeze-thaw-freeze cycle is also necessary. For example, serum and plasma can be apportioned into 200 μL, DNA and RNA can be apportioned into 75 μg and 1 μg, respectively.

**Fresh frozen tissue.** OCT-protected frozen tissue should be cryosectioned for hematoxylin and eosin staining to assess tumor area and percentage as part of the annotation. The information will guide case selection as well as steps in macro and micro dissection. Integrity of the sample must be maintained through sterile processes to avoid degradation and/or contamination. The tissue must be kept frozen until submerged into extraction buffer. There are many choices of methods for the extraction of cellular components such as RNA, DNA, and proteins.

**Fixed tissue.** Tissue biopsies and/or resection must undergo complete and adequate fixation in a buffered environment. This is important for the morphologic review as well as the evaluation of reference and novel markers. Neutral buffered formalin has been widely used and is most appropriate for current and future translational processes. As previously mentioned, there are many choices of methods for the extraction and isolation of DNA; however, there are only few methods for the extraction of RNA and protein from fixed tissue. Care must be taken in making selections to avoid loss of valuable tissue resource.

Sectioning should be performed with new blades and clean water. Nonprecipitating formalin should be avoided.
Xylene rather than its substitutes should be used.
Manual rather than automated processes should be used for small, starting samples.
Washes between steps in any and all processes must be complete to avoid reagent carry over.
CURRENT NOVEL PROCESSING OF BLOOD AND FIXED TISSUE

Isolation of Mononuclear Cells

Monocytes, a type of white blood cells, are found both in bone marrow and blood. The following procedure describes the isolation of monocytes in peripheral blood (peripheral blood mononuclear cells, PBMC).

Steps

1. Centrifuge CPT tubes containing whole blood for 15 minutes at 25°C at 2700 rpm.
2. Pipet out the top layer (plasma) and transfer into two labeled tubes and freeze.
3. Pipet out the lymphocyte layer (the cloudy layer) with a transfer pipette, placing the layer into a fresh 15-mL tube. Fill tube with room-temperature PBS.
4. Centrifuge for 10 minutes at 1800 rpm.
5. Discard supernatant and resuspend pellet in 1 mL PBS.
6. Fill tube with PBS, inverting eight times to mix and wash.
7. Centrifuge for 15 minutes at 2000 rpm.
8. Discard supernatant.
9. Proceed to cell sorting or resuspend and freeze pellet in 1 mL of RPMI or DMSO (dimethyl sulfoxide) for later sort.

Note: For blood collected in any other anticoagulated tube, dilute 1:1 with RPMI, layer on Histopaque 1077, then proceed with step 1.

Specialized Sorting of Circulating Tumor/Endothelial Cells

1. Magnetic labeling (~45 min)
   a. Centrifuge sample for 10 minutes at 300g at room temperature.
   b. Carefully remove the supernatant completely.
   c. Resuspend cell pellet in appropriate amount of buffer as specified by the microbeads package insert. For fewer cells, use the same volume.
   d. Add microbeads per package insert.
   e. Incubate for 15 minutes at 6°C to 12°C. For less than $5 \times 10^6$ total cells, use the same volume.
   f. Wash cells by adding 10 to 20 times the labeling volume of buffer and centrifuge at 300g for 10 minutes.
   g. Remove supernatant completely (save the supernatant in appropriately labeled 1 mL cryovial) and resuspend cell pellet in appropriate amount of buffer (1 mL of buffer/10^8 total cells).
   h. Proceed to magnetic separation.
2. Magnetic separation (~60 min)
   a. Turn on the manual and/or automated cell sorter.
   b. When the machine has finished initializing, perform a column exchange to remove blank columns and insert magnetic columns. (To save time, try to perform these functions while the sample and beads are incubating.)
   c. Wipe down probes with alcohol to disinfect.
   d. Place the sample in the uptake position, and place empty 15-mL conical tubes in the rack under the Neg1, Pos1, and Pos2 probes.
   e. Choose “Separation” on the main menu.
   f. Use the down arrows to select “possel_s” and press “start.”
   g. For both the positive and negative cells, separate elution container must be used (~15 min).
   h. Prepare and/or submit samples for next step in processing (DNA extraction, thin prep slides, etc.).

**Extraction of RNA from Peripheral Blood in PAXgene**

1. Centrifuge the PAXgene Blood RNA Tube for 10 minutes at 3000g to 5000g using a swing-out rotor.
2. Remove the supernatant by decanting or pipetting. Add 4 mL RNase-free water to the pellet, and close the tube using a fresh secondary Hemogard closure.
3. Vortex until the pellet is visibly dissolved, and centrifuge for 10 minutes at 3000g to 5000g using a swing-out rotor. Remove and discard the entire supernatant.
4. Add 350 μL Buffer BR1, and vortex until the pellet is visibly dissolved.
5. Pipet the sample into a 1.5-mL microcentrifuge tube. Add 300 μL Buffer BR2 and 40 μL proteinase K. Mix by vortexing for 5 seconds, and incubate for 10 minutes at 55°C using a shaker/incubator at 400 to 1400 rpm. After incubation, set the temperature of the shaker/incubator to 65°C (for step 20).
6. Pipet the lysate directly into a PAXgene Shredder spin column (lilac) placed in a 2-mL processing tube, and centrifuge for 3 minutes at maximum speed (but not to exceed 20,000g).
7. Carefully transfer the entire supernatant of the flow-through fraction to a fresh 1.5-mL microcentrifuge tube without disturbing the pellet in the processing tube.
8. Add 350 μL ethanol (96–100%, purity grade p.a.). Mix by vortexing, and centrifuge briefly (1–2 seconds at 500–1000g) to remove drops from the inside of the tube lid.
9. Pipet 700 μL sample into the PAXgene RNA spin column (red) placed in a 2-mL processing tube and centrifuge for 1 minute at 8000 to 20,000g.
10. Pipet the remaining sample into the PAXgene RNA spin column, and centrifuge for 1 minute at 8000 to 20,000g. Place the spin column in a new 2-mL processing tube, and discard the old processing tube containing flow-through.

11. Pipet 350 µL Buffer BR3 into the PAXgene RNA spin column. Centrifuge for 1 minute at 8000 to 20,000g. Place the spin column in a new 2-mL processing tube and discard the old processing tube containing flow-through.

12. Add 10 µL DNase I stock solution to 70 µL Buffer RDD in a 1.5-mL microcentrifuge tube. Mix by gently “flicking” the tube and centrifuge briefly to collect residual liquid from the sides of the tube.

13. Pipet the DNase I incubation mix (80 µL) directly onto the PAXgene RNA spin column membrane, and place on the bench top (20–30°C.) for 15 minutes.

14. Pipet 350 µL Buffer BR3 into the PAXgene RNA spin column, and centrifuge for 1 minute at 8000g to 20,000g. Place the spin column in a new 2-mL processing tube and discard the old processing tube containing flow-through.

15. Pipet 500 µL Buffer BR4 to the PAXgene RNA spin column, and centrifuge for 1 minute at 8000g to 20,000g. Place the spin column in a new 2-mL processing tube, and discard the old processing tube containing flow-through.

16. Add another 500 µL Buffer BR4 to the PAXgene RNA spin column. Centrifuge for 3 minutes at 8000g to 20,000g.

17. Discard the tube containing the flow-through, and place the PAXgene RNA spin column in a new 2-mL processing tube. Centrifuge for 1 minute at 8000g to 20,000g.

18. Discard the tube containing the flow-through. Place the PAXgene RNA spin column in a 1.5-mL microcentrifuge tube, and pipet 20 µL Buffer RNAs-free water directly onto the PAXgene RNA spin column membrane. Centrifuge for 1 minute at 8000g to 20,000g to elute the RNA.

19. Repeat the elution step by adding the eluted RNA to the same PAXgene RNA spin column membrane. Centrifuge for 1 minute at 8000g to 20,000g to elute the RNA.

20. Incubate the eluate for 5 minutes at 65°C. in the shaker/incubator without shaking. After incubation, chill immediately on ice.

21. If the RNA samples will not be used immediately, store at −20°C or −70°C.

Laser Capture Microdissection

Laser capture microdissection (LCM) allows for the microscopic removal of specific cells from paraffin or frozen tissue sections mounted on glass slides.
This procedure describes the isolation of a pure population of cells followed by its nucleic acid or protein extraction.

1.1. Prepare paraffin section on glass slide.
   a. Collect 5 μm paraffin section onto uncoated glass slide, using microtome and water bath.
   b. Allow to air-dry overnight at room temperature or at 42°C for up to 8 hours.
   c. Deparaffinize and lightly stain slide(s) by immersing in the following:
      i. Xylene × 2: 5 minutes each
      ii. 100% ethanol: 1 minute
      iii. 95% ethanol: 1 minute
      iv. 70% ethanol: 1 minute
      v. Distilled water: 1 minute
      vi. Harris hematoxylin: 1 dip (Fisher’s Protocol #245–678)
      vii. Tap water: 2 minutes
      viii. Distilled water: 1 minute
      ix. 95% ethanol: 1 minute
      x. Eosin Y: 1 dip (Fisher’s Protocol #314–631)
      xi. 95% ethanol: 1 minute
      xii. Allow to air-dry

1.2. Prepare frozen section on glass slide.
   a. Collect 8 μm frozen section on an uncoated glass slide, using cryostat.
   b. Store frozen until ready to postfix.
   c. Postfix and lightly stain.
   d. Remove a maximum of four slides from cryostat or freezer, place on clean paper towel, i.e., KimWipe, and allow to thaw for approximately 30 seconds.
   e. Place the slides in 75% ethanol for 30 seconds.
   f. Transfer the slides to distilled water for 30 seconds.
   g. Place slides on KimWipe, apply 100 μL of histogene staining solution, and stain for approximately 20 seconds.
   h. Place the slides in distilled water for 30 seconds.
   i. Place the slides for 30 seconds in each of the following ethanols in increasing concentration: 75%, then 95%, and then 100%.
   j. Transfer the slides to xylene for 5 minutes.
   k. Place the slides on a KimWipe to dry in a ventilation hood for 5 minutes.

2. Dissect area of tissue from slide (to be performed same day as steps 1.1 or 1.2).
   a. Place slide on Pixcell II LCM [Arcturus (now part of Molecular Devices), Mountain View, California, U.S.] and locate area of interest using 4× or 10× objective.
b. Load CapSure HS caps.
c. Turn on laser key, push laser enable button, and locate laser spot on screen.
d. Focus laser beam to bright, well-defined spot.
e. Test laser, observing melted plastic ring(s).

See following table for power and duration.

<table>
<thead>
<tr>
<th>Laser spot size (μm)</th>
<th>Power (mW)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>CapSure HS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small (7.5)</td>
<td>65–75</td>
<td>650–750 μs</td>
</tr>
<tr>
<td>Medium (15)</td>
<td>35–45</td>
<td>2.5–3.0 ms</td>
</tr>
<tr>
<td>Large (30)</td>
<td>45–55</td>
<td>6.0–7.0 ms</td>
</tr>
</tbody>
</table>

f. View captured cells on cap.
g. Move cap to capping station.
h. Transfer to ExtractSure device from CapSure HS cap.

3. Extract DNA from dissected tissue.

a. Pipet 155 μL of reconstitution buffer into one vial of proteinase K. Gently vortex the tube to mix the reagents and immediately place the tube on ice. Completely dissolve reagent, excessive mixing can denature proteinase K.
b. Assemble the ExtractSure Sample Extraction Device onto the CapSure HS LCM Cap.
c. Using clean forceps, remove the cap from the cap insertion tool and place into the alignment tray (sample facing up).
d. Make sure that the cap snaps securely and lies flat on the bottom of the opening of the alignment tray.
e. Position the ExtractSure Device over the cap using clean forceps. The fill-port should be facing up.
f. Using forceps, push the ExtractSure Device down onto the cap until it snaps securely into place. Use forceps to ensure that the cap is firmly attached to the ExtractSure Device.
g. Pipet 10 μL of proteinase K extraction solution into the microchamber formed by the assembled ExtractSure Device and CapSure HS LCM Cap.
h. Using gloved hands, place a 0.5-mL thin-walled reaction tube over the ExtractSure device.
i. Cover the assembly in the alignment tray with the incubation block and incubate it overnight at 65°C.
j. After incubation, remove the assembled CapSure HS LCM Cap and ExtractSure Device from the incubator, place the assembly into a microcentrifuge and centrifuge for 1 minute at 4000g.
k. Separate the CapSure HS LCM Cap and ExtractSure Device. Close the microcentrifuge tube containing the extract and heat to 95°C for 10 minutes in a heating block to inactivate the proteinase K. Cool the sample to room temperature.

**Isolation of DNA, RNA, Protein Using TRIZOL Method**

**Isolation of RNA by TRI Reagent**

*Homogenization*  Homogenize fresh tissue samples (that have been stored in RNAlater) in TRI Reagent (1 mL/50–100 mg tissue) using a glass-Teflon or Polytron homogenizer. Sample volume should not exceed 10% of the volume of TRI Reagent used for homogenization.

*Phase separation*

1. Store the homogenate for 5 minutes at room temperature to permit the complete dissociation of nucleoprotein complexes.
2. Add 0.2 mL chloroform to the homogenate, cover the samples tightly, and shake vigorously for 15 seconds.
3. Store the resulting mixture at room temperature for 2 to 15 minutes and centrifuge at 12,000 g for 15 minutes at 4°C.
4. Following centrifugation, the mixture separates into a lower red phenol-chloroform phase, interphase, and the colorless upper aqueous phase. RNA remains exclusively in the aqueous phase, whereas DNA and proteins are in the interphase and organic phase. The volume of the aqueous phase is about 60% of the volume of TRI Reagent used for homogenization.

*RNA precipitation*

1. Transfer the aqueous phase to a fresh tube.
2. Precipitate RNA from the aqueous phase by mixing with 0.5 mL of isopropanol.
3. Store samples at room temperature for 5 to 10 minutes and centrifuge at 12,000 g for eight minutes at 4°C to 25°C. RNA precipitate (often invisible before centrifugation) forms a gel-like or white pellet on the side and bottom of the tube.

*RNA wash*

1. Remove the supernatant and wash the RNA pellet (by vortexing) with 75% ethanol (1 mL).
2. Centrifuge at 7500 g for 5 minutes at 4°C to 25°C.
**RNA solubilization**

1. Remove the ethanol wash and briefly air-dry the RNA pellet for 3 to 5 minutes.
2. Dissolve RNA water by passing the solution a few times through a pipette tip and incubating for 10 to 15 minutes at 55 to 60°C.

**Isolation of DNA by TRI Reagent**

The DNA is isolated from the interphase and phenol phase separated from the initial homogenate as described in the RNA isolation protocol.

**DNA precipitation**

1. Remove the remaining aqueous phase overlying the interphase. Precipitate DNA from the interphase and organic phase with ethanol.
2. Add 0.3 mL of 100% ethanol and mix samples by inversion.
3. Store the samples at room temperature for 2 to 3 minutes and sediment DNA by centrifugation at 2000 g for 5 minutes at 4°C to 25°C.

**DNA wash**

1. Wash the DNA pellet twice in a solution containing 0.1 M trisodium citrate in 10% ethanol (1 mL).
2. Store the DNA pellet in the washing solution for 30 minutes at room temperature with periodic mixing and centrifuge at 2000g for 5 minutes at 4°C to 25°C.
3. Suspend the DNA pellet in 75% ethanol (1.5–2 mL of 75% ethanol). Store for 10 to 20 minutes at room temperature with periodic mixing and centrifuge at 2000g for 5 minutes at 4°C to 25°C. This ethanol wash removes pinkish color from the DNA pellet.

**DNA solubilization**

1. Remove the ethanol wash and briefly air-dry the DNA pellet by keeping tubes open for 3 to 5 minutes at room temperature.
2. Dissolve the DNA pellet in 8 mM NaOH by slowly passing through a pipette. Add (0.3 mL) of 8 mM NaOH.
3. Centrifuge at 12,000g for 10 minutes and transfer the resulting supernatant containing DNA to new tubes.

**Isolation of Proteins by TRI Reagent**

**Protein precipitation**

1. Aliquot a portion of the phenol-ethanol supernatant (0.2–0.5 mL, 1 volume) into a microfuge tube.
2. Precipitate proteins by adding 3 volumes of acetone.
3. Mix by inversion for 10 to 15 seconds to obtain a homogeneous solution.
4. Store samples for 10 minutes at room temperature and sediment the protein precipitate at 12,000g for 10 minutes at 4°C.

**Protein wash**
1. Decant the phenol-ethanol supernatant and disperse the protein pellet in 0.5 mL of 0.3 M guanidine hydrochloride in 95% ethanol + 2.5% glycerol.
2. Disperse the pellet using a pipette tip. After dispersing the pellet, add another 0.5 mL aliquot of the guanidine hydrochloride/ethanol/glycerol wash solution to the sample and store for 10 minutes at RT.
3. Sediment the protein at 8000g for 5 minutes.
4. Decant the wash solution and perform two more washes in 1 mL each of the guanidine/ethanol/glycerol wash solution. Disperse the pellet by vortexing after each wash to efficiently remove residual phenol.
5. Wash in 1 mL of ethanol containing 2.5% glycerol.
6. Sediment the protein at 8000g for 5 minutes.
7. Decant the alcohol, invert the tube, and dry the pellet for 7 to 10 minutes at room temperature.

**Protein solubilization**
1. After briefly air-drying the protein pellet, add 1% SDS (0.2 mL) to the protein pellet.
2. Gently disperse and solubilize the pellet for 15 to 20 minutes by flicking the tube or pipetting as required.

**Storage of RNA, DNA, and/or Protein**
Store the solubilized proteins and DNA at −20°C or colder, store RNA at −70°C or colder.

**Tissue Microarray Creation**
This procedure describes the use of the MTA-1 (Manual Tissue Arrayer) by Beecher Instruments, Inc. (Sun Prairie, Wisconsin, U.S.) to create tissue microarray (TMA) blocks. TMA is a technology which facilitates research by permitting the collection of biological specimens representing a great many individuals without requiring a great deal of storage space. TMA eases comparison of the histological characteristics of different patients’ pathologies by placing many different tissue samples together on the same microscope slide for observation.
Tissue is the Issue

Also, TMA helps prevent the unnecessary exhaustion of irreplaceable pathological material, since it permits the same diagnostic and research objectives to be achieved with far less tissue than was required for such objectives before TMA existed.

A core of paraffin is removed from a “recipient” paraffin block (one embedded without tissue) and the remaining empty space is filled with a core of paraffin-embedded tissue from a “donor” block. This process can be repeated up to 230 times for one TMA block, resulting in a block containing mapped pieces from many different blocks, and when sectioned, produces a slide containing mapped pieces from many different blocks. A corresponding map records each core’s location in the grid and the information connected with it (patient ID number, study arm, etc.). The array sections can be used for all histological staining, including hematoxylin & eosin (H&E), IHC, and ISH.

1. Prepare the donor blocks and corresponding H&E slides.
   a. Re-embed tissue, as needed, for more uniform wax consistency and optimal tissue orientation.
   b. Donor tissue should be at least 1.0 mm thick for best results.
   c. Prepare a fresh H&E slide from each re-embedded potential donor block.
2. Pathologist characterizes each H&E slide for appropriateness of diagnosis and percentage of tumor.
   a. Examine each H&E slide microscopically and indicate, with marker, the area of interest (AOI). This area corresponds to the area on the block from which the core(s) will be taken.
3. Prepare the recipient (TMA) block.
   a. Label a blank cassette, using the automated cassette labeler.
   b. Determine the size of the TMA block and use:
      M—the medium mold (30 mm × 23 mm) or
      L—the large mold (35 mm × 23 mm)
   c. Make a paraffin block, containing no tissue, using the labeled cassette.
   d. Assure that the paraffin block surface is flat and parallel to the underside of plastic cassette and that it contains no air bubbles.
4. Design the array.
   a. Determine the diameter of the cores to be removed from the donor blocks. Available core sizes include 0.6, 1.0, 1.5, and 2.0 mm. Core size is often indicated by protocol, i.e., breast cancer studies require the smallest core size—0.6 mm—to conserve tissue while removing three cores from each donor block during the routine production of three TMA blocks from each group of subjects.
   b. Allow for enough space (3.0 mm) around the edge of the block to avoid cracking paraffin.
c. Determine the number of columns and number of rows to use by noting the following table, which describes the maximum number of cores per TMA block, with respect to core size and mold size.

<table>
<thead>
<tr>
<th>Core size (mm)</th>
<th>Spacing between cores (mm)</th>
<th>Mold size</th>
<th>Maximum no. of cores/spaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6&lt;sup&gt;a&lt;/sup&gt; yellow</td>
<td>1.5</td>
<td>LM</td>
<td>20 × 12 = 240</td>
</tr>
<tr>
<td>1.0&lt;sup&gt;a&lt;/sup&gt; green</td>
<td>2.0</td>
<td>LM</td>
<td>15 × 9 = 135</td>
</tr>
<tr>
<td>1.5&lt;sup&gt;a&lt;/sup&gt; black</td>
<td>2.5</td>
<td>LM</td>
<td>12 × 7 = 84</td>
</tr>
<tr>
<td>2.0&lt;sup&gt;a&lt;/sup&gt; white</td>
<td>3.0</td>
<td>LM</td>
<td>10 × 6 = 60</td>
</tr>
</tbody>
</table>

<sup>a</sup>Punches are color coded.

5. Create a map, identifying each core’s position on the grid its donor block.
   a. Enter information into Excel Program (TMA Mapping Tool, Attachment 1) located on Server/Pathcore/TMA. TMA may be mapped using alternate methods.
   b. Include cores from tissue and cell line control blocks, in duplicate.
      i. Tissue blocks, i.e., for breast protocols:
         - Fibroadenoma
         - Normal endometrium
         - Normal salivary gland
         - Normal testis
      ii. Cell line blocks, i.e., for breast protocols:
         - MCF7
         - SKBR3
         - T47D
   c. Each TMA block will be asymmetric to assure map orientation, i.e.:
      - 2nd column blank,
      - 4th row blank, and
      - bottom right space blank

6. Construct the TMA.
   a. Install the pair of tissue punches, for desired core size, in MTA-1. Install the slightly larger of the pair (identified in blue) on the right (for the donor block). Install the slightly smaller of the pair (identified in red) on the left for the (for the recipient or TMA block).
   b. Insert the recipient (TMA) paraffin block into the MTA-1 block holder. Secure in place by tightening two screws in block holder with MTA-1 allen wrench.
c. Remove the first core from TMA block.
   i. Assure that recipient punch is swung into place.
   ii. Align punch over site of first core, usually top left corner, by
       turning micrometers.
   iii. Zero micrometers by pressing the X and Y “ZERO/ABS” buttons.
   iv. Gently lower punch into TMA block, using care not to punch too
       deeply. Maximum punch depth can be adjusted with large vertical
       screw on the left.
   v. Raise punch.
   vi. Remove paraffin core from punch, using punch’s stylet.

d. Remove core from first donor block.
   i. Place the donor block bridge over the TMA block holder.
   ii. Swing donor punch into place.
   iii. Manually align donor block so that AOI is directly under the
        donor punch.
   iv. Lower and raise punch.
   v. Remove core of tissue from punch using punch stylet.

e. Place donor core into TMA block.
   i. Remove donor block bridge from over the TMA block holder,
      noting location of delicate loose core of tissue.
   ii. Using forceps, align donor core over hole created for it in the
       TMA block.
   iii. Lower donor core into hole and assure flat level block surface by
        pressing evenly with clean glass slide.

f. Adjust the micrometer(s) to move the precision guide to the next X/Y
   position. Follow the Spacing Between Cores (mm) Column in the
   maximum number of cores per TMA block table above to determine
   the appropriate increments.

g. Repeat this cycle (steps 6c–6f) to construct the whole array.

h. To punch holes on right half of TMA block, rotate it 180° and replace
   in recipient block holder, because the x axis micrometer should not be
   extended beyond 9 mm.

i. Smooth and level surface of TMA block.
   i. Remove the TMA block from the recipient block holder.
   ii. Place the TMA block in oven at 37°C, maximum temperature, for
       30 minutes.
   iii. Gently press a glass slide on the top of the TMA block, applying
       even pressure to push all tissue cores on the block to same level.
7. Section TMA block.
   i. Cut a section from the TMA block and stain the slide with H&E for QA/QC review by the pathologist.
   ii. Store TMA block and save TMA map for any future sectioning requests.
       Label side of block
       2nd column left blank.
       4th row left blank
       Bottom right corner left blank
Acquisition and Preservation of Tissue for Microarray Analysis

Brian R. Untch and John A. Olson
Institute for Genome Sciences and Policy and the Department of Surgery, Duke University Medical Center, Durham, North Carolina, U.S.A.

INTRODUCTION

Increasingly, collection of normal and neoplastic human tissue is being incorporated in the design of clinical trials and in routine clinical settings under tissue banking programs. The goal behind these activities is either to facilitate identification and development of robust predictive or prognostic biomarkers or to enable basic investigation of tumor or host biology.

As experience in tissue banking has broadened, the role of rigorous collection protocols, tissue handling, and annotation has become increasingly appreciated. To this end, several groups have published guiding principles and standard operating procedures (SOPs) designed to advance the fields of tissue banking, biomarker development, and tumor biology (1,2).

This chapter is designed to summarize key points in the process of informed consent, tissue collection, tissue preservation and storage, as well as extraction of macromolecules of interest.

INFORMED CONSENT

Tissue banking requires informed consent by patients donating their tissues. The informed consent document should include standard topics such as study design, study purpose, and patient risks/benefits. Privacy rights are paramount for
patients when their health information is stored in databases, perhaps indefinitely. The Health Insurance Portability and Accountability Act and the Common Rule are federal regulations that require protection of patient identification. Database information and patient identifiers (name, social security number, institutional medical record ID) should remain separate. An arbitrary code system should be created to uniquely label specimens and data. The informed consent document should specify which investigator will keep the anonymous identifiers and patient identifiers together. In the event that third-party sharing of patient identifier data becomes necessary, written authorization must be obtained from the patient. Also specified in the informed consent document is the patients’ right to withdraw from the study or have their specimen removed from the tissue databank. Contact information for these requests is provided to the patient at the time of document signing.

A recent topic of debate has been whether giving consent for tissue banking also allows for indiscriminate investigations related to and unrelated to the initial study design. Implicit to the potential of tissue banking is the use of specimens for areas of research that are unanticipated or currently unknown. However, the use of specimens for research not specifically noted in the informed consent document may violate religious and personal beliefs. The inclusion of broad consent for any possible project has been proposed by some as a solution (3). Others argue that this does not constitute informed consent, as the possibilities for potential research are unpredictable (4,5). Our approach has been to specifically describe research to be performed on collected specimens. This prevents any misunderstanding between patient and researcher. This is not useful for large repositories that collect tissue for future investigations.

TISSUE ACQUISITION

When tissues are devascularized and removed from the body, cellular metabolism is altered as hypoxia and acidification occur and nutrient delivery ceases. At a minimum, macromolecules in ischemic tissues are subject to degradation. Further, cellular response to ischemia may lead to stress-specific changes in gene expression. This may ultimately lead to changes in gene and protein expression or degradation until the cellular machinery stops or tissue preservation occurs. To accurately and reproducibly analyze gene and protein expression events within cells, tissue must be preserved as quickly and uniformly as possible.

For all protocols where banking is superimposed upon clinical care, above all else, banking strategies must not interfere with pathologic assessment of either the tumor or specimen resection margins. When designing these protocols, involvement of the pathologist who is custodian of the tissue record is critical. This is not to say that banking cannot occur before specimen processing in pathology, but that overall specimen integrity must be maintained during the research specimen collection (e.g., core biopsy of lumpectomy specimens). When a pathologist is immediately available, specimen banking after pathologic
Acquisition and Preservation of Tissue for Microarray Analysis

gross evaluation is preferred. However, if a pathologist is not immediately available, then tumor specimens must be collected before pathologic assessment but without specimen disruption. Regardless of when during the processing of the pathologic specimen banking occurs, representative samples of all banked tissue should be reviewed by the pathologist to both ensure that critical diagnostic material is not lost and to document that research tissue samples are indeed tumor tissues.

To further ensure that research tissue acquisition will not have an impact on diagnosis, frozen specimens should be held and not analyzed until definitive pathologic assessment of the entire remaining specimen has been completed. At the time of definitive diagnosis of the surgical specimen, research specimens can then be deposited in the tumor bank (Fig. 1). If analysis of the surgical specimen does not demonstrate the abnormal pathology, then research specimens may be

---

**Figure 1** Flow diagram demonstrating the sample handling of core biopsy specimens in patients suspected of having breast cancer. Research and clinical core biopsies should be pathologically examined so that the patient’s diagnosis is not affected and that research specimens contain the tissue of interest.
used for frozen-section analysis. Should the frozen section of research specimens reveal abnormal histology not present in the surgical specimen (e.g., neoplasia is present in the research specimens and not in the surgical specimen), all samples can be turned over to pathology for routine processing. While this may result in a short delay as surgical pathology is reevaluating frozen research specimens, this can be explained in a preoperative consent form.

In all cases, frozen research specimens should be held frozen until two weeks after a definitive pathologic report is issued on the diagnostic specimens. This hold period is desirable, as once DNA and RNA are extracted, no further diagnostic information relevant to patient care will be available.

For clinical trials and research protocols, specimen collection needs to be standardized so that tumor samples are treated in a similar manner to minimize technical variability in gene and protein expression data. Several cooperative groups have developed banking SOPs for collection of formalin-fixed specimens and blood samples; however, few have until recently focused on frozen-tissue collection. The American College of Surgeons Oncology Group (ACOSOG) has developed effective procedures for collecting tumor tissue from breast cancer patients in neoadjuvant studies with good success. The European Human Frozen Tumor Tissue Bank has also developed similar SPOs that are followed across institutions and countries in an effort to reduce variability (2). In the former case, specialized collection kits have been privately developed and are in use by ACOSOG investigators for multisite tissue and blood collection at the time of diagnostic biopsy and again at surgical resection (Fig. 2). Technicians responsible for tumor collection prepare reagents and containers before excision.

*Figure 2* A stereotactic biopsy device to retrieve breast cancer samples from excised lumpectomy specimens. (A, arrow) Excised lumpectomy specimens are placed in the device in an oriented fashion, and a specimen radiogram is taken to identify the location of the tumor. (B) Coordinates are taken from the radiogram and are used to guide core biopsies taken from the tumor within an example lumpectomy specimen held in fixed orientation in the device.
Samples are required to be snap frozen within five minutes of excision, regardless of the method of tissue fixation (described subsequently). Rapid and uniform handling of specimens from patient to preservative is essential.

TECHNIQUES OF CRYOPRESERVATION

RNA degradation by tissue RNases can be detrimental to RNA yield and quality, with implications on microarray and quantitative PCR experiments. Traditional practice has been to snap freeze samples in liquid nitrogen followed by permanent storage at a minimum of $-80^\circ C$. This prevents changes to the RNA, and tumor banks using this method have been highly successful in producing high-quality extracted RNA for microarray analysis. Given the complexity of sample acquisition and multi-institutional collection, recent studies have tried to determine the importance of immediate ex vivo freezing. Micke et al. (6) used ex vivo tumors at room temperature to determine the amount of time until RNA degradation occurred. Distinct ribosomal peaks were identified up to 16 hours, indicating that this tissue can remain stable for extended periods of time. Other sample-handling issues, such as freeze-thawing cycles, have also been studied. Jochumsen et al. (7) established that gene expression analysis is not altered in up to three freeze-thaw cycles. Despite these reports, it has been our experience that RNA degradation can be extremely variable, even when SOPs are employed (Figs. 3,4). To ensure quality RNA, samples should be frozen as soon as possible after being devascularized.
Preservation of tissue morphology can be difficult in samples that are snap frozen. These samples are also subject to desiccation or freezer burn. Samples can be immersed in optimal cutting temperature (OCT) compound to preserve tissue morphology and prevent sample damage. This method has been shown to generate reproducible genomic DNA extraction results (8). Our group has not observed differences in microarray data when using OCT (unpublished observation). However, Turbett et al. (9) reported inhibition of a polymerase chain reaction (PCR) assay when compared with non-OCT compound–preserved specimens. OCT compound can also make specimen thawing and dissection more difficult. When specimens contain a mix of tissue (as in the case of undissected tumor specimens with margins), separating whole tumor for RNA extraction can be difficult, particularly when trying to avoid time-dependant RNA degradation. Nonetheless, few studies have investigated OCT compound and its potential to alter microarray results. A recommendation for or against its use cannot be made at this time.

Because of the challenges of sample handling (pathology analysis) and the potential for RNA degradation, RNA stabilization buffers have been developed. Commercially available RNAlater (Ambion inc.) has been shown to maintain RNA quality before snap freezing. Several studies have shown consistent microarray results when storing samples in RNA stabilization buffers (10–12). One such study found that samples could be stored up to one hour at 37°C, without affecting the RNA or subsequent microarray data (13). These buffers
will be important for tissue acquisition protocols, particularly with regard to multi-institutional trials that will require transferring tissue specimens to a central tissue bank. This approach allows the samples to be shipped refrigerated rather than frozen. The principal disadvantages are that proteins are degraded, and subsequent sectioning for histologic verification of tumor identity is problematic.

PARAFFIN-EMBEDDED TISSUE FOR MICROARRAY ANALYSIS

The discovery of paraffin wax is attributed to the noted hydrocarbon chemist, Dr. Carl (Karl) Ludwig von Reichenbach (1788–1869). After being developed in 1830, pathology laboratories discovered its potential as a method for storing and preserving specimens. Throughout the twentieth century, formalin-fixed, paraffin-embedded tissue became a reliable way to archive tissue. Given the large quantity of specimens available in paraffin, these tissue banks were identified as a possible tissue source for array-based techniques. Numerous challenges have been identified that complicate the macromolecular extraction process from these samples.

Sample handling is an important step that can lead to variability observed in microarray results. Ischemia time can profoundly affect protein, DNA, and RNA extraction because of enzymatic degradation. As noted above, sample collection is ideally performed by snap-freezing tissue in liquid nitrogen immediately after being devascularized and removed from a patient. Sample handling of formalin-fixed, paraffin-embedded tissue is variable, and this makes application of the array data potentially unreliable.

Many investigators have attempted to develop methods for extraction of these samples with mixed results. Given the cross-linking effects of formalin, it was thought that protein extraction for proteomic analysis would be impossible. However, attempts have been made to isolate proteins from paraffin, but results have been mixed and more technical development is required (14).

Extraction of nucleic acids can be accomplished using commercially available proteinase digestion buffers (15). This method also eliminates the risk of Rnase activity in the sample as all previously fixed proteins are destroyed. Formalin fixation does not appear to affect the yield of RNA, unless the tissue is fixed for an extended period of time. This can result in cross-linking between RNA and protein (16). As long as the methylol groups added to the bases during fixation can be removed, the quality of the RNA is not perturbed (17) and results have been comparable with frozen-tissue controls (18).

The most difficult problem encountered with these specimens has been that the total DNA or RNA yield from these samples is inadequate and often degraded (19). Yield of DNA or RNA is dependent on duration of protein digestion, temperature, and pH (20). Most array platforms will require several micrograms of RNA for hybridization. Amplification of the RNA has been suggested as means to increase quantity, but these techniques can result in considerable variability between samples (21). Further, amplification of partially
degraded samples can also result in unreliable data. Statistical adjustments for nucleic acid amplification in microarray experiments are being developed and may eventually prove to be useful.

As extraction, amplification, and statistical techniques improve, formalin-fixed, paraffin-embedded tissue may provide an additional resource for researchers. Nevertheless, the variability in sample handling history makes application of the array data from these samples less than reliable. At this point in time, use of these samples for application to clinical investigations is discouraged.

FINE NEEDLE ASPIRATION, CORE NEEDLE BIOPSY, AND SURGICAL SPECIMENS IN MICROARRAY ANALYSIS

Using microarray analysis, numerous genomic signatures have been developed that attempt to predict tumor phenotypes, including lymph node metastasis, tumor growth patterns, postoperative margin positivity, chemosensitivity, and oncogenic pathway activation. Most of these analyses have been based on samples collected at the time of surgery. These large volume samples can yield adequate RNA for array hybridization and also afford the opportunity to assess multiple areas of the tumor and subcellular compartments via laser capture microdissection. However, tumor size and growth is known to affect gene expression, and successful application of the developed models to earlier stage, preoperative biopsy specimens may not be possible.

Nonetheless, application of these predictive tests depends on obtaining adequate tissue for microarray analysis before a definitive surgical intervention. The preoperative diagnosis of breast cancer is typically achieved though two methods of tumor biopsy: core needle biopsy (CNB) and fine needle aspiration (FNA). For microarray-based predictive testing, an optimal protocol would allow these two modalities to collect tissue for pathologic analysis and for molecular extraction.

FNA is performed by inserting a small gauge needle, typically under image guidance, into a tumor or lesion. Tissue is removed through a negative-pressure syringe. Tumor architecture is typically destroyed during this process, but the resultant sample can be analyzed under the microscope for tumor cell characteristics. Likewise, this small amount of tissue can be harvested for RNA and microarray analysis. However, concern exists over adequate sample amount (22). Several studies have shown that FNA can produce tissue in amounts appropriate for pathology review and array applications (23,24).

CNB is performed under image or visual guidance in a fashion similar to FNA. However, with CNB, a larger bore needle is used, and the resultant specimen has maintained tissue architecture, allowing for a more accurate pathologic analysis. This larger amount of tissue can also be used for array-based studies. Several studies have addressed adequacy of tissue acquisition in core
biopsy samples (13,24). Studies from our laboratory indicate that core biopsy specimens can be collected before surgery and can undergo successful RNA extraction (unpublished observation). The average yield from core biopsy specimens was 10.4 μg of RNA, enough for most array platforms. After extraction, RNA samples underwent hybridization to the HU95 Affymetrix microarray platform and then chemosenstivity and oncogenic pathway signatures were successfully applied. This is consistent with the work of Ellis et al. (13), who was able to extract RNA from core biopsy specimens and distinguish cancer and noncancer phenotypes.

LASER CAPTURE MICRODISSECTION

Recently, focus has shifted away from whole-tumor microarray analysis to analysis of individual cell types within a tumor. A tumor is composed of a heterogeneous group of cells, malignant cells, premalignant cells, normal parenchymal cells, and supportive cells. This cellular heterogeneity must be taken into account when applying array-based techniques (25). Numerous successful expression-profiling studies have been performed using whole-tumor macromolecular extraction (26–28). These analyses thus incorporate a variety of different cell types. Laser capture microdissection allows one to isolate histologically distinct cell types and perform subsequent analysis. The technique employs an infrared laser in conjunction with light microscopy. Prepared histologic slides are covered with a transfer film. When cells of interest are identified, the laser is activated over the cells. This process causes the cells to attach to the transfer film. The remainder of the sample remains, and the isolated cells can undergo extraction procedures. Notably, this procedure can result in lower yields of cells requiring amplification of nucleic acids for use in microarray experiments (29,30). Laser capture microdissection has the potential to not only improve our understanding of cellular deregulation but also to help determine how other cells within the tumor contribute to the overall tumor phenotype. Immune cells, vascular structures, and normal parenchymal cells within a tumor may contribute significantly to the growth and metastatic potential. Expression analysis of these individual cell types may help elucidate their roles.

SUMMARY

Acquisition and preservation of tissue for microarray experiments requires understanding of informed consent, sample handling, macromolecular extraction, and sampling methods. Significant advances in these areas have been made in recent years to improve the quality of microarray data. Appreciating these aspects is paramount if the application of microarray techniques is to contribute meaningfully to the care of the cancer patient.
REFERENCES

**Acquisition and Preservation of Tissue for Microarray Analysis**


Tapping into the Formalin-Fixed Paraffin-Embedded (FFPE) Tissue Gold Mine for Individualization of Breast Cancer Treatment

Mark Abramovitz
VM Institute of Research, Montreal, Québec, Canada

Brian Leyland-Jones
Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia, U.S.A.

INTRODUCTION

It is now clearly recognized that breast cancer is a heterogeneous disease that requires personalized targeted therapy in order to have the greatest impact on survival. This necessitates the development of biomarkers that can be used to predict outcome and response to therapy. Unfortunately, very few biomarkers have made their way from the research bench to the clinic. The main reason for this is that in order to validate breast tumor markers, large sample sets from well-defined patient groups are required, although this could be overcome by tapping into large numbers of well-defined banked tumor specimens. The efforts that are being made, though, toward the development of biomarkers in order to improve the assessment of cancer prognosis and diagnosis are great, given the enormous potential that they embody.
Formalin-fixed, paraffin-embedded (FFPE) tissue samples represent the largest collection of well-annotated, clinical tumor samples that are readily available for conducting retrospective studies and, therefore, represent a potential gold mine from which important biomarkers are just waiting to be discovered, given that the right tools can be brought to bear upon this invaluable resource. However, until relatively recently, a major caveat regarding the use of these FFPE specimens for genomewide studies was that due to the formalin fixation process, the standard procedure for handling tumor tissue samples in pathology labs around the world, RNA degradation occurred, greatly limiting their potential utility. In fact, with regard to nucleic acid–based assays, FFPE specimens have been mainly subjected to reverse transcribed-polymerase chain reaction (RT-PCR) analysis. This chapter will describe the progress that has been made in greatly expanding the use of FFPE tumor specimens for several genomic-based applications that will greatly increase our ability to tap into this gold mine and extract the biomarker and genomic data nuggets that will improve cancer prognosis, diagnosis, and treatment.

EXTRACTION OF RNA FROM FFPE TUMOR TISSUE SAMPLES FOR GENE EXPRESSION ANALYSIS

FFPE tissue samples, a vast archive of pathologically well-characterized clinical samples from randomized trials, are an enormous potential resource that will allow translational scientists to more fully describe breast cancers at the molecular level and will ultimately have important etiologic and clinical implications. Even though the degradation of RNA that occurs because of the formalin fixation process results in an RNA species with an average size of approximately 200 nt (1), several groups have illustrated that it is feasible to extract and purify RNA from such fixed tissue and perform gene expression profiling (2–9). With the development of real-time quantitative (q) RT-PCR technology that has high detection sensitivity and high dynamic range, it is now possible to detect even rare messages in FFPE tissue and to examine the variation of expression over quite a large dynamic range. Amplicons are designed specifically on small segments of DNA [<100 base pairs (bp)] to achieve close to 100% efficiency for all amplicons, regardless of their length and nucleic acid composition. The potential of this technology becomes even more attractive in that it permits the analysis of thousands of tissue samples available through existing banks and without the need to collect the relatively complicated, freshly collected frozen tissue. Some studies have addressed improvements to the process of isolating high-quality FFPE RNA suitable for qRT-PCR or high-throughput gene expression profiling (10–13).

Several groups have exploited this technology to evaluate and validate prognostic and predictive markers of human disease, the response to certain cancer therapies and to predict recurrence. Initial work by Cronin et al. (2004) (14) from Genomic Health, Inc., which led to the development of the Oncotype DX clinical assay for breast cancer (see below), reported developing an RT-PCR
method with multianalyte capability for potential use in clinical research and diagnostic assays. Two multianalyte assays, one with 48 genes and the other with 92 genes, were designed and tested in different experiments with preserved human breast cancer tissue. In the first case, all of the genes profiled in FFPE RNA yielded measurable values, and the overall profile was validated by its similarity to that generated with well-preserved RNA from matched freshly frozen tissue. In the second experiment using a 92-gene assay, only one of the tested genes failed to yield a signal. Measured levels of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) mRNAs were concordant with the levels of the respective proteins as measured by immunohistochemistry (IHC) at an independent clinical reference laboratory. Approximately 90% concordance was obtained when RT-PCR expression results for ER and PR were dichotomized into positive and negative values and compared with ER- and PR-positive and ER- and PR-negative assignments based on IHC. It is noteworthy that similar levels of concordance were found in other comparative studies of IHC versus RT-PCR (15,16).

DEVELOPMENT OF AN RT-PCR-BASED BREAST CANCER CLINICAL DIAGNOSTIC TEST: ONCOTYPE DX ASSAY

A major development in this area has been the development of a multigene qRT-PCR assay to predict the likelihood of breast cancer recurrence in node-negative, ER-positive, tamoxifen-treated patients (17). The authors studied the likelihood of distant recurrence in breast cancer patients who had involved lymph nodes and ER-positive tumors that were poorly defined by clinical and histopathologic measures. They tested whether a prospectively-defined 21-gene RT-PCR assay of FFPE tumor tissue could predict the likelihood of distant recurrence in patients in the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14 clinical trial. Measured levels of gene expression for 16 cancer-related and 5 reference genes were used, by a prospectively defined algorithm, to calculate a recurrence score (RS) and risk group (low, intermediate, or high) for each patient. Adequate RT-PCR profiles were obtained in 668 of 675 tumor blocks. The proportion of patients categorized as low, intermediate, or high risk by the RT-PCR assay was 51%, 23%, and 27%, respectively. The Kaplan–Meier estimates and 95% confidence intervals for the rates of distant recurrence at 10 years were 6.8% (4.0%, 9.6%), 14.3% (8.3%, 20.3%), and 30.5% (23.6%, 37.4%), respectively, for the low-, intermediate-, and high-risk groups; the rate for the low-risk group was significantly lower than the rate for the high-risk group ($p < 0.001$). In a multivariate Cox model, RS provides significant ($p < 0.001$) predictive power that goes beyond age and tumor size. The RS was also predictive of overall survival ($p < 0.001$) and could be used as a continuous function to predict distant recurrence for individual patients. Finally, the RS was prospectively validated to predict the likelihood of distant recurrence in tamoxifen-treated breast cancer patients with node-negative, ER-positive tumors (18,19).
This methodology has been sufficiently successful, so that the Food and Drug Administration (FDA) approved it in 2004 to become commercially available [Genomic Health, Inc. (GHI), Redwood City, California, U.S.]. Oncotype DX is a diagnostic assay that quantifies the likelihood of breast cancer recurrence in women with newly diagnosed stage I or II, node-negative, ER-positive breast cancer who will be treated with tamoxifen. More importantly, this test will pave the way for individualization of cancer therapy. While the St. Gallen classification distinguished 7.9% of patients as low risk, the GHI RS identified 50.6% of patients as low risk. Similarly, St. Gallen overestimated the high-risk patients at 58.8%, while the GHI test classified high-risk patients at only 27.1%. It is obvious that St. Gallen underestimated the low-risk patients and overestimated the high-risk ones. The advantage of the newly developed multigene RS is, therefore, that low-risk patients can be saved from potentially toxic, high-dose chemotherapy.

RNA-BASED EXPRESSION PROFILING OF FFPE BREAST TUMOR SPECIMENS

Expression Profiling: The DASL Assay

Gene expression profiling has become an important aspect in prediction, prognosis, and cancer modeling. Gene signatures resulting from expression profiling studies have the potential to define cancer subtypes, predict the clinical outcome (recurrence of disease) and response to specific therapies, and analyze oncogenic pathways (20–23). Investigations of gene pathways and interactions indicated by gene signatures that are truly predictive of the clinical endpoints are necessary to understand the biology underlying this predictive value. More importantly, when combined with clinical and demographic factors, multiple forms of molecular (protein- and gene-based) data can provide information that has the potential to identify unique characteristics of individuals and so lead to individualized treatment strategies (24,25).

Although most array-based assays utilize RNA prepared from frozen specimens, the newer cDNA-mediated annealing, selection, extension, and ligation (DASL) assay (Illumina, Inc., San Diego, California, U.S.) platform was specifically designed to profile small transcripts typically found in FFPE tissue (26–31). The DASL platform is based on massively multiplexed RT-PCR applied in a microarray format that allows for the determination of expression of up to 512 genes (502 genes in the standard cancer panel available from Illumina) using a minimal amount of total RNA (100 to ~200 ng per assay) isolated from 96 FFPE tumor tissue samples in a high-throughput format (29,30). In the procedure, biotinylated random nonamers (biotin-d(N)9) and oligo d(T)18 are used for cDNA synthesis, and therefore probes are designed so that they can target any unique region of the gene without limiting the selection of the optimal probe to the 3’ ends of transcripts. In addition, because of the small size of the targeted gene sequence (50 nucleotides), along with the use of random primers in
the cDNA synthesis, RNAs that are otherwise too degraded for conventional microarray analysis can be readily detected. Sequence-specific query oligonucleotides encompassing primer extension, ligation, and universal PCR in highly-plexed reactions (1536-plex), two-color labeling, and redundant (~30-fold redundancy of each bead type) feature representations are used to probe up to three different sites on each cDNA (Fig. 1A) and, therefore, lend the assay the necessary sensitivity and reproducibility for quantitative detection of differential expression using RNA from FFPE tissue (29,30). In an initial study, a 231-gene
cancer panel was used in the DASL assay in order to profile both breast and colon cancer FFPE tumor samples. Cluster analysis was able to separate breast from colon tissue types and, subsequently, divide each tissue sample set into cancer versus normal (32). More recently, the DASL assay has been used to identify RNA signatures in prostate cancer, including a 16-gene set that correlates with prostate cancer relapse (31,33).

We have designed our own custom 512-gene breast cancer–related gene panel so that it incorporates previously identified signature genes from various breast cancer microarray studies that have been used in the classification of breast tumors (34,35), in prognosis (MammaPrint) (36) and as predictors of outcome to treatment (17,37–40), as well as a host of additional genes that have been implicated as playing a role in a number of breast cancer–related processes, including proliferation, angiogenesis, (41,42), metastasis (43,44), DNA repair, apoptosis (45), thrombosis (46), and custom-selected genes taken from the most recently published data on breast cancer (47–49). Additional cancer-related genes include oncogenes, cell cycle genes, telomerase-related genes, amplified genes, breast cancer stem cell genes, and senescence-related genes (50–52).

**MiRNA Profiling**

MiRNAs (miRNAs) are small noncoding RNA species that regulate the translation and stability of mRNA messages for hundreds of downstream target genes via partial complementarity to short sequences in the 3′-untranslated region (3′UTR) of mRNAs. The aberrant expression or mutation of miRNAs, a distinct class of RNAs that control gene expression via translation repression or degradation, has been noted in cancers and suggests a key role for miRNAs in tumorogenesis. MiRNAs can act both as tumor suppressors and as oncogenes, depending on the sets of downstream targets that they regulate (53). Studies have recently implicated nearly 30 of the over 450 known human miRNAs in breast cancer, as tumor suppressor genes or as oncogenes, with significant deregulation among only a handful (54). Others are correlated with specific tumor types as identified by protein expression. For example, the mir-30 cluster is downregulated in both ER-negative, PR-positive and ER-positive, PR-negative tumors (54). Differential expression has also been noted by tumor stage or proliferative indices. The tumor suppressor, p53, was recently shown to regulate expression of mir-34b and mir-34c (55). One key set of tumor suppressor miRNAs includes mir15a-16, which targets the translation of the suppressor of apoptosis, Bcl-2 (56). These observations suggest that specific miRNAs are associated with major breast cancer subgroups, including hormone-positive and hormone-negative subgroups and the basal or triple negative subtype, and that the deregulated miRNAs may differ between African-American and Caucasian women.

These low-molecular weight miRNAs can be extracted at the same time as total RNA is being isolated from FFPE tissue samples. qRT-PCR can be used to screen tumor samples for miRNA expression. The TaqMan™ array human
miRNA panel from Applied Biosystems allows for the expression profiling of 365 miRNAs from one RNA sample. Illumina, Inc., has recently adapted its DASL assay for high-throughput multiplexed miRNA expression profiling (Fig. 1B). It has made a human miRNA panel that can be used to detect 470 miRNAs described in the miRBase database and an additional 265 potential miRNAs that have been identified in an RNA-primed array-based Klenow extension (RAKE) analysis study (57,58). With these arrays and tools now available, our understanding of the role that miRNAs play in breast cancer will be greatly enhanced.

Whole-Genome Expression Profiling

In order to conduct expression profiling using microarrays, it has been necessary to obtain freshly frozen tissue as the starting material for RNA extraction. This presents limitations with regard to obtaining enough samples, enough tissue per sample, as well as the logistics of storing this type of material. Most recently, new technologies have come into play that enable the use of RNA extracted from FFPE to be used in genomewide gene expression profiling. The company NuGEN Technologies, Inc. (San Carlos, California, U.S.) has developed an amplification system that can take picogram amounts of total RNA to yield micrograms of amplified cDNA that can be fragmented and labeled for same-day hybridization to a variety of array formats, including Affymetrix, Agilent, and Illumina microarrays (59). This type of technology will undoubtedly generate significant microarray data from FFPE breast tumor samples in the near future further adding to the utility of FFPE specimens.

DNA-BASED ARRAY PROFILING THAT EMPLOYS FFPE SAMPLES

Array CGH and SNP Profiling

DNA copy number alterations (CNAs), which are the result of genomic instability, can give rise to gene amplifications and gene deletions that play important roles in tumor progression, and it has become apparent that an increased number of CNAs correlate with a poor prognosis for the patient. Several genes that undergo CNAs have been identified that act as either oncogenes or tumor suppressors, which are crucial factors in the disease process. With regard to breast cancer, array comparative genomic hybridization (CGH) has been used principally to identify novel breast cancer-related genes, genes that can be used to classify different types of breast cancer, genes that yield prognostic or predictive information, or genes that may represent important therapeutic targets. Single nucleotide polymorphisms (SNPs) have been linked to individual variations in breast cancer susceptibility. Genetic variations in genes involved in ER and drug metabolism and genes involved in cell cycle control as well as variations depending on ethnic background have been found to contribute to breast cancer susceptibility.
Array CGH has been used in the classification of breast tumors and has shown that subtypes can be discerned on the basis of genomic instability patterns. In one study, array CGH, made up of 2464 genomic clones, was applied to the analysis of CNAs in 62 sporadic invasive ductal carcinomas, and three breast tumor subtypes were delineated based on genomic DNA CNAs (60). In another study, 89 breast cancer tumors of the luminal A, luminal B, HER2, and basal-like subtypes were also found to yield distinct CNA profiles, suggesting that genomic instability mechanisms differ between subtypes (61). Most array CGH studies have made use of DNA prepared from fresh or frozen tumor samples; however, the ability to tap into archival FFPE clinical samples would greatly accelerate the rate at which disease-related genomic change, including translocation breakpoints, could be more easily identified. Recent studies have addressed this important issue and shown that DNA can be prepared from FFPE tissue samples and used successfully in array CGH (62–64). Van Beers and colleagues (2006) have also developed a multiplex PCR predictor that can be used to assess the quality of the DNA extracted from FFPE samples (65) prior to use in array CGH. Recently, Oosting and co-workers (2007) (66) have compared the Affymetrix GeneChip and Bac array CGH with the Illumina BeadArray platform, which is a high-density SNP microarray that allows for the measurement of both genomic copy number and loss of heterozygosity (LOH). The BeadArray platform (Fig. 2) is amenable for use with fragmented DNA from FFPE tissue because of the fact that only short, intact genomic segments of approximately 40 bp flank each SNP on the array. The authors found that when paired comparisons of DNA prepared from freshly frozen and FFPE tissue were conducted, the patterns of genomic aberrations were basically identical.

The use of array CGH and SNP profiling in conjunction with the vast archive of FFPE clinical breast cancer samples will allow for the identification of chromosomal aberrations as well as novel variants that contribute to breast cancer, and this will lead to uncovering novel breast cancer-related genes that will be key to understanding the molecular mechanisms underlying the disease.

**Methylation Profiling**

Epigenetic changes, which involve very subtle molecular modifications, can cause activation or suppression of genes that can lead to abnormalities resulting in tumorigenesis (67), and these epigenetic changes have been found to occur in many cancers, including breast, colon, prostate, and blood cancer. The epigenotype is affected by such factors as the environment, age, epigenotype of the parents, and the genotype at the loci that regulate DNA methylation and chromatin, and these factors can cause aberrant methylation of the cytosine residue of DNA, resulting in gene regulation.

Illumina has developed an array-based platform (Fig. 3) that can be used in conjunction with its standard Cancer Methylation Panel, which spans 1505 CpG loci selected from 807 genes with 71% of those selected genes containing at least
2 CpG sites. The Cancer Methylation Panel includes genes selected from various functional classes, including oncogenes and tumor suppressor genes; genes involved in cell cycle control, DNA repair, differentiation, and apoptosis; and X-linked and imprinted genes (68). Although DNA prepared from FFPE tumor samples has not yet been specifically validated for use in the assay, it would be useful to be able to scan large numbers of FFPE samples for methylation patterns, which could provide important clues regarding tumor type.

THE WAY FORWARD

The goal is to obtain as complete a molecular picture of a breast tumor as possible, which can now be realistically undertaken using FFPE tissue specimens, truly untapped gold now ready to be brought to the surface and subjected to a thorough nucleic acid characterization. Profiling platforms can now be brought to bear on these invaluable samples in order to build a comprehensive picture of the tumor. In the scheme shown in Figure 4, FFPE samples, either sections or cores, are used to extract DNA, RNA, and miRNA. These nucleic acids can now be
Figure 3 The DNA methylation assay. (A) In order to perform a DNA methylation assay, genomic DNA must first undergo bisulfite conversion. Cytosines are converted to uracils, but methyl-cytosine is unreactive. (B) In order to take advantage of this differential reactivity to bisulfite treatment, two pairs of complex oligonucleotide probe sets are designed for each CpG site. The first probe pair is made up of an ASO and a LSO to interrogate the methylated state of the CpG site and a corresponding ASO-LSO pair for the unmethylated state. The ASO is used to determine whether a site is methylated or not and also incorporates a universal PCR primer sequence, P1 or P2. P1 and P2 are fluorescently labeled, each with a different dye, and associated with the “T” (unmethylated) allele or the “C” (methylated) allele, respectively. The LSO is made up of three parts: a CpG (Continued)
locus-specific sequence, an address sequence in the middle corresponding to a complementary address sequence on the bead array, and a universal PCR priming site (P3) at the 3'-end. The pooled assay oligonucleotides are first annealed to bisulfite-converted genomic DNA. An allele-specific primer extension step is then carried out using ASOs that are extended only if their 3'-base is complementary to their cognate CpG site in the genomic DNA template. Allele-specific extension is followed by ligation of the extended ASOs to their corresponding LSOs to create PCR templates. Common primers P1, P2, and P3 are then used to amplify the ligated products, which are subsequently hybridized to a microarray bearing the complementary address sequences. **Abbreviations:** ASO, allele-specific oligonucleotide; LSO, locus-specific oligonucleotide; PCR, polymerase chain reaction.
subjected to a panoply of analysis platforms that can be used to detect CNAs, analyze SNPs, and investigate methylation aberrations, regulated gene expression, and signaling pathway changes, as well as miRNA alterations. Data derived from tumor tissue analyses can be used to further our understanding of the different subtypes, which could lead to new insights into breast cancer etiology, prediction of disease outcome, response to therapy, and ultimately to better and personalized patient management.

REFERENCES

The Formalin-Fixed Paraffin-Embedded (FFPE) Tissue Gold Mine

INTRODUCTION

The quest for optimal breast cancer therapy has been thwarted by recurrent treatment failures. As a result, the necessity has emerged to concomitantly rethink the fundamental basis of breast cancer, and to ensure that it is adequately and optimally targeted by therapeutic agents, and combinations of these agents. Adding to the complexity of this disease is the heterogeneity of breast cancer, and the realization that breast cancer is not a single disease, but a family of diseases. The daunting challenge that we face with a newly diagnosed breast cancer patient is one-to-one communication that tailors an individualized treatment strategy.
based on the corresponding breast cancer pathophysiology. At the same time, we are mandated to address the urgency of this therapy on the basis of the underlying clinical assessment, the risks of metastases and death, as well as the benefits and adverse effects associated with therapeutic regimens to ensure that an enlightened decision among available therapeutic options can be mutually reached with an individual patient. As our understanding of the pathophysiology of breast cancer is incomplete, it is critical to channel the information that we glean from observations that deviate from our current models into a unified paradigm shift where therapeutic targets are validated, and subsequently attacked and destroyed. Given the complexity of breast cancer, the highest level of therapeutic success is most likely to be achieved when the uniqueness of an individual patient’s breast cancer is matched with individualized, combined, and targeted therapy.

CANCER STEM CELL HYPOTHESIS

The cancer stem cell hypothesis is a major player in the current paradigm shift in oncology (1,2). Cancer is, in essence, a genetic disease where three groups of genes are responsible for tumor progression, namely oncogenes, tumor suppressor genes, and stability genes (3,4). Cancer is a multistage process involving the accumulation of genetic and epigenetic changes in genes involved in the regulation of cell growth, DNA repair, and metastasis. In the traditional stochastic model, the vast majority of cancer cells can proliferate extensively to form new tumors. In contrast, the cancer stem cell model suggests that only rare cancer stem cells have the intrinsic ability to proliferate extensively to form new tumors (Fig. 1) (1). Clinical observations indicate that breast cancer responses to therapy do not necessarily correlate with patient survival. Tumor-initiating cells with the concomitant stem cell properties of self-renewal and successive maturation of progenitor cells into highly differentiated cells are referred to as “cancer stem cells” (Fig. 2).

The cancer stem cell hypothesis provides an explanation for our clinical observations, including those which deviate from the stochastic model, by suggesting that therapies that kill cancer cells but not cancer stem cells may lead to tumor regression, but that the disease will subsequently recur as the cancer stem cells act as lethal seeds to repopulate the tumor (5). Thus, tumor regression in itself may represent an incomplete clinical end point, as it reflects cell death limited to differentiated tumor cells, while sparing cancer stem cells (Fig. 3). The “dandelion hypothesis” has been used to illustrate this pattern of clinical activity as it is analogous to cutting off a dandelion (or other weed) at ground level, which may initially appear to produce the desired effect, but only the elimination of the root will prevent the weed from regrowing (6). Therefore, effective therapies should target cancer stem cells, with limited effects on normal cells. In the presence of therapies that kill cancer stem cells, tumors are predicted to lose their ability to generate new cancer cells, subsequently degenerate, leading to complete tumor eradication.

Stem cells are multipotent with high proliferative potential and with concomitant properties of self-renewal and differentiation into multilinage
cells. Recent findings suggest that stem cell properties are integral and fundamental to the formation and perpetuation of human cancers, including breast cancer (7). The quiescence, self-renewal, and differentiation of stem cells are regulated by both intrinsic and extrinsic mechanisms. Intrinsic mechanisms include those governing the epigenetic state of stem cells, for example, hematopoietic stem cells are controlled by chromatin remodelers such as polycomb group proteins (8,9). The extrinsic mechanisms include changes in stem cell fate that are dictated by the environment, i.e., the stem cell niche. As our tissues are organized into stem cell hierarchies ranging from stem cells with extensive proliferative and self-renewal capacity to mature cells with little or no capacity for cell division, it is likely that cancer cell populations might also be organized in stem cell hierarchies, ranging from a small number of cells that are responsible for fueling the uncontrolled growth of the tumor cells to a large array of differentiated daughter cells (10). While the vast majority of cells in a tumor are merely the nontumorigenic daughter cells of cancer stem cells, the highly proliferative capacities of stem cells may drive the continued expansion of malignant cells. As cellular proliferation decreases with differentiation, the long-lived and slowly dividing stem and progenitor cells are likely candidates for the accumulation of mutations associated with carcinogenesis (11,12).
The multilineage differentiation of cancer stem cells may contribute to the heterogeneity of certain tumors, including breast cancer. The parallels between normal stem cells and cancer stem cells suggest that some cancer stem cells arise from mutated normal stem cells, and are responsible for the proliferative capacity of cancers (Fig. 2) (7). The observation that stem and cancer cells share certain signaling pathways that regulate their self-renewal suggests that normal stem cells can, in fact, give rise to cancer cells (8). Mutated progenitor cells which acquire self-renewal properties may also become cancer stem cells (Fig. 2) (7,13). Epigenetic gene silencing and associated promoter CpG island DNA hypermethylation may precede genetic changes in premalignant cells and foster the accumulation of additional genetic and epigenetic hits that contribute...
Breast Cancer Stem Cells and Their Niche: Lethal Seeds in Lethal Soil

Figure 3 Therapies which kill cancer cells, but not CSCs lead to an initial tumor regression, but as the CSCs are not eliminated, the tumor will regenerate. In contrast, therapies that target CSCs also eliminate the genesis of new tumor cells and result in tumor regression. Abbreviation: CSCs, cancer stem cells.

to the generation of cancer stem cells from stem or early progenitor cells (14). The accumulation of mutations in cancer stem cells may disrupt the tight control of normal stem cells, leading to the dysregulation of self-renewal, tumorigenesis, loss of asymmetric division, and aberrant differentiation (Fig. 2) (12,15). The fate of stem cells, including self-renewal and differentiation, may also be programmed by the stem cell niche, i.e., the cellular microenvironment providing necessary support and stimuli to sustain self-renewal and other environmental factors (11,13,16–19). In normal breast development, mammary stem cells differentiate into myoepithelial, alveolar epithelial, and ductal epithelial cells (20). Cancer stem cells may acquire additional features associated with tumor progression, metastases, and therapeutic failure, including genetic instability and drug resistance (11). Mechanisms of therapeutic resistance include cellular quiescence and the acquisition of protein expression responsible for the cellular efflux of chemotherapeutic agents. The elimination of the cancer stem cell compartment of a tumor may be necessary to achieve stable, long-lasting cancer remissions. Therapeutic strategies designed to eradicate the cancer stem
compartment should selectively, or at least preferentially, target the self-renewal, survival, and proliferative pathways specific to cancer stem cells, or induce differentiation. Thus, advances in our knowledge of stem cell properties are pivotal to the specific and effective targeting of these cells in therapeutic strategies, including cancer therapies.

CURRENT STATUS OF BREAST STEM CELL RESEARCH

The human mammary gland is organized during development as a hierarchy of stem and progenitor cells that are successively limited in their multilineage potential and proliferative capacities. Stem cells represent approximately 1 in 250 epithelial cells of the normal human breast (17). The transplantation of mammary cells into cleared murine fat pads is a functional assay used to identify mammary epithelial cells with stem cell properties based on their ability to generate tumors (21,22). The existence of a stem cell origin of murine mammary development was established using transplantation of cells at limiting dilutions. These experiments confirmed the clonal nature of the regenerated alveolar, ductal, lobular, and more complex structures (22–24). The unequivocal demonstration of the existence of a murine mammary stem cell occurred with the reconstitution of a complete and functional mammary gland from a single lineage negative (Lin−)/cluster of differentiation (CD)24+/CD29+ cell, marked with a β-galactosidase (LacZ) transgene (25).

Three types of murine mammary epithelial cell progenitors have been identified thus far (22–24). The first is a bipotent progenitor with features of both luminal and myoepithelial characteristics, the second has luminal features, and the third has myoepithelial features. Analogous human studies of stem and progenitor cells have been limited by the challenges associated with the development of suitable in vivo xenotransplantation assays. Alternative strategies have been undertaken to assay human breast stem cells and to optimize conditions that support the in vitro growth and differentiation of primitive human epithelial cells. These include the proliferation of nonadherent mammospheres in suspension from human breast cancer tissues (17,26) and murine mammary fat pads (27). These mammospheres are enriched in stem and progenitor cells, as illustrated by their ability to differentiate along all three mammary epithelial lineages and to clonally generate complex functional structures in reconstituted three-dimensional culture systems (17). Additional studies to characterize stem and progenitor cells include the identification of side populations using flow cytometry, i.e., low Hoechst 33342 dye-retaining populations (28–30). Side populations are enriched 30-fold in mammospheres derived from human tissues (17).

The use of mammosphere assays subsequently lead to the discovery that the Notch signaling pathway plays a pivotal role in normal human mammary development through its activity in both stem and progenitor cells, concurrently affecting self-renewal and lineage-specific differentiation. As a result, it was proposed that abnormal Notch signaling contributes to mammary carcinogenesis.
by dysregulating the self-renewal of normal mammary stem cells (26). Additional pathways thought to be involved in stem cell self-renewal include hedgehog (HH), Bmi-1, wingless (Wnt/β-catenin) and phosphatase and tensin homolog deleted on chromosome ten (PTEN) (11,20,31).

Further studies in human breast stem/progenitor cells identified a subset of human breast cancer cells using flow cytometry with the phenotype CD44+/CD24+/Lin− that possess highly tumorigenic properties in a nonobese diabetic/severe combined immunodeficiency (NOD/SCID) xenograft model (21). As few as 200 cells with the epithelial-specific antigen (ESA)+, CD44+/CD24−/Lin− phenotype consistently generated tumors in mice, whereas even 20,000 cells with alternative phenotypes were ineffective. The isolation and serial passage of the tumorigenic CD44+/CD24−/Lin− cell population in NOD/SCID mice gave rise to additional tumorigenic CD44+/CD24+/Lin− cells, as well as phenotypically diverse nontumorigenic cells, that composed the bulk of the tumors, thereby confirming the self-renewal and differentiating capacities of these cells. While the unequivocal demonstration of stem cell characteristics necessitates model systems capable of tumor generation from a single cell, a population of CD44+/CD24−/Lin− cells appears to exhibit properties of human breast stem/progenitor cells in NOD/SCID mice.

**GENETIC SIGNATURES AND POLYMORPHISMS WHICH MAY INFLUENCE BREAST CANCER DEVELOPMENT AND TREATMENT OUTCOMES**

The study of molecular breast cancer portraits identifies patterns of molecular breast cancer markers using gene expression microarrays, which subsequently are useful in the selection of therapy. These portraits also seek to overcome potential differences in drug sensitivity or target expression between tumor-initiating cells and the more frequent nontumorigenic cells (32). The rationale for the pursuit of new drug discovery based on molecular targets results from the significant increment of the effectiveness of these drugs as compared with conventional chemotherapy. A recent systematic analysis of the consensus of human breast and colorectal cancers included the analysis of 13,023 genes in 11 breast and 11 colorectal cancers, and revealed that individual tumors accumulate an average of approximately 90 mutated genes, but that only a subset of these contribute to the neoplastic process (33). Using stringent criteria, 189 genes (average of 11 per tumor) were found to be mutated at significant frequencies. The vast majority of these genes were not known to be genetically altered in tumors and are predicted to affect a wide range of cellular functions, including transcription, adhesion, and invasion. The genetic landscape of breast and colorectal cancer defined in this study provides new targets for diagnostic and therapeutic intervention.

Gene signatures of tumor invasiveness in tumorigenic CD44+/CD24−/low breast cancer cell populations correlate with overall and metastasis-free survival,
independently of established clinical and pathological variables (34). CD44^+CD24^-/low breast cancer cells were also shown to increase from a median of 4.8% to 14.8% after chemotherapy in paired breast cancer biopsies from 35 patients obtained before and after 12 weeks of neoadjuvant chemotherapy. Microarray analyses of CD44^+CD24^-/low breast cancer cells using the Affymetrix U133 GeneChip demonstrated an increased expression of self-renewal pathways. An increase in mammosphere formation of residual tumors was also observed postchemotherapy. Taken together, these results suggest that therapy resistant cancer stem cells may exist in the neoadjuvant setting of breast cancer therapy, and that breast cancer patients may benefit from the therapeutic targeting designed to eradicate these cells (35).

In addition to chemoresistance, progenitor cells may also be resistant to radiation. This radioresistance has been shown in an animal model to be mediated at least in part by Wnt signaling, which has been implicated in stem cell survival (36). Radioresistance was investigated by treating primary BALB/c mouse mammary epithelial cells with clinically relevant doses of radiation. An enrichment in normal progenitor cells [stem cell antigen (SCA) 1-positive and side population progenitors] was identified.

Specifically, radiation selectively enriched for progenitors in mammary epithelial cells isolated from transgenic mice with activated Wnt/beta-catenin signaling but not for background-matched controls. Irradiated stem cell antigen 1-positive cells had a selective increase in active beta-catenin and survivin expression compared with SCA 1-negative cells. Radiation also induced enrichment of side population progenitors cells of the MCF-7 human breast adenocarcinoma cell line. In comparison to differentiated cells, progenitor cells have different cell survival properties that may facilitate the development of targeted antiprogenitor cell therapies.

Closely intertwined with response prediction are an individual patient’s genetic polymorphisms, which may increase breast cancer susceptibility and influence the efficacy or safety of breast cancer treatments. The accumulation of genetic factors that underlie breast cancer susceptibility or therapeutic escape may trigger the transformation of a normal stem or progenitor cell into a cancer stem cell. The variation of treatment responses between individuals also lies in differences in DNA sequences that alter the expression or function of proteins that are drug targets. These proteins are encoded by genes identified in early studies that are linked to highly penetrant, single-gene traits. Germline polymorphisms which confer susceptibility to breast cancer are associated with the highly penetrant breast cancer genes BRCA1 and BRCA2, p53, PTEN, ataxia telangiectasia mutated (ATM), and candidate BRCA3 genes (37). In addition, BRCA1 has been suggested to function as a breast stem cell regulator (38). Thus, BRCA1 may contribute to a conceptual link between hereditary and sporadic breast cancer. The loss of BRCA1 function may produce a block in cell differentiation and an increase in the self-renewal of mammary stem cells, resulting in the expansion of the stem cell pool. Taken together with its known function in
DNA repair, BRCA1 may generate genetically unstable stem cells which are targets for additional carcinogenic events and the aggressive triple negative “basal” stem cell–like phenotype (39).

Low-penetrance genes that may influence breast cancer susceptibility alone or in combination with environmental factors include genes in DNA repair, metabolic pathways, and estrogen-related pathways including HRAS1 minisatellite, CHEK2, XRCC1, XRCC2, RAD51, XRCC3, and LIG4 (37). Metabolic enzymes such as N-acetyltransferases (NAT) and glutathione-S-transferases (GST) are involved in the metabolic inactivation of carcinogens. Mutations in these enzymes can result in increased toxicity and exposure to mutagenic substances, thereby increasing mutations, which contribute to carcinogenesis. Superoxide dismutases (SODs) are metalloenzymes that protect cells from free radicals, and their mutations can lead to increased breast cancer susceptibility (37). Cytochrome p450s (CYP) metabolically activate carcinogens into chemically reactive electrophiles. CYP1B1 is the most common form in the breast, and its alanine and serine (ALA/SER) polymorphism significantly increases breast cancer susceptibility (37). Additional targets of polymorphisms that increase breast cancer susceptibility include COMT, CYP17, CYP19, PPAR-γ, TGF-β, HSP70-1, HSP70-2, and HSP70-HOM (37). Germline polymorphisms may also influence an individual’s susceptibility to the effects of breast cancer treatments (37). For example, dihydropyrimidine dehydrogenase (DPD) and thymidylate synthase (TS) are pivotal to the metabolism of 5-FU, and specific polymorphisms of these enzymes lead to 5-FU toxicity. Paclitaxel is metabolized by cytochrome P450, and specific CYP polymorphisms lead to defective metabolism. Additional breast cancer treatments with known polymorphisms, which may effect drug efficacy or safety include tamoxifen (ER, CYP2D6, SULT1A1), aromatase inhibitors (CYP19, CYP1A2, CYP2C9, CYP3A), cyclophosphamide (GST), methotrexate (MTHFR), doxorubicin (GST), and epirubicin (UGT2B7) (40). Future advances hinge on the more difficult challenge of elucidating multigene determinants of breast cancer susceptibility and drug responses through an intersection of genomics and medicine, which has the potential to yield a new set of molecular diagnostic tools that can be used to individualize and optimize drug therapy (41). Additional refinements can be accomplished through coupling in vitro drug sensitivity data with microarray data, to develop gene expression signatures that predict sensitivity to individual chemotherapeutic drugs (42). Thus, as we continue to fine-tune our stratification, we also approach our ultimate goal of defining individual molecular breast cancer portraits.

**STEM CELL SANCTUARIES AS A MECHANISM OF THERAPEUTIC RESISTANCE**

Targeting of stem cells is particularly complex since this cellular population is largely quiescent. Thus, therapeutic agents directed against cycling cells are predictably ineffective in this population, which is essentially shielded from
these drugs. Cancers that respond to therapy initially may acquire drug resistance during the course of treatment, while other cancers appear to be intrinsically resistant. The cancer stem cell hypothesis suggests that in both of these cases, the resting cancer stem cell, which is both the cancer-initiating cell and its source of replenishment under selective pressure, has innate drug resistance by virtue of its quiescent stem cell phenotype. Acquired drug resistance in more differentiated cancer cells may occur through gene amplification or rearrangement, thereby contributing to an aggressive phenotype. Through the accumulation of mutations that distinguish cancer stem cells from their normal counterparts, these cells may acquire additional features associated with tumor progression, metastases, and therapeutic failure, including genetic instability, radioresistance, and chemoresistance (11,36). In the case of gliomas, cancer stem cells contribute to radioresistance through the preferential activation of the DNA damage checkpoint response and an increase in DNA repair capacity (43). Radioresistance may also result from a selective increase of cell survival pathways in progenitor cells, including β-catenin and survivin (36).

Stem cell niches regulate stem cell proliferation and cell fate decisions, and they also play a protective role by shielding stem cells from environmental insults. In this manner, vascular niches might protect brain cancer stem cells from chemo- and radiotherapies, and the opportunity for these cells to reform a tumor mass in spite of an initial clinical response (44). With the acquisition of properties of tumor progression and metastases, cancer stem cells may travel to sanctuary sites, such as across the blood-brain barrier, where it may be difficult to eradicate them, even with targeted therapies.

The drug resistance of normal tissue stem cells is constitutively mediated by multidrug resistance (MDR) transporters and detoxifying enzymes (45). DNA repair mechanisms, resistance to apoptosis, and telomerase activity also contribute to the stability of normal tissue stem cells. Among the mechanisms responsible for stem cell drug resistance, there is mounting speculation that ABC transporters repress the maturation and differentiation of stem cells (46,47). However, since breast cancer resistance protein 1 (BCRP1) knockout mice display normal hematopoiesis instead of accelerated hematopoietic maturation, it is clear that multiple factors determine the maturation and differentiation of stem cells (48). In addition, protection from hypoxia appears to be another function of ABC transporters. As a result, it has been postulated that Bcrp expression protects hematopoietic stem cells from hypoxic environments by preventing porphyrin accumulation that causes mitochondrial death (49).

Thus, leading edge-targeted breast cancer therapies must overcome all of the obstacles associated with cancer stem cell biology, including both the intrinsic and extrinsic properties of resistance, which define these enigmatic cells and the cloisters that keep them safe. From the cancer stem cell hypothesis, it follows that systemic administration of an efficacious MDR reversal agent would render tumor and normal tissue stem cells equally susceptible to the chemotherapeutic agents, offering no net gain in therapeutic index (45). Thus, targeting
the properties that distinguish cancer stem cells from their normal stem cell counterparts is necessary.

CELLULAR BREAST CANCER STEM CELL TARGETS

The identification of cancer stem cell markers is necessary for the isolation and purification of cancer stem cells as they are potential targets of novel drug discovery and development. The functional targets that define stem cells are their abilities to self-renew and differentiate using serial passages in xenotransplantation models. The current breast cancer stem cell hypothesis suggests that the most primitive human breast stem cell does not express the estrogen receptor (ER) while more mature progenitors may or may not express ER depending on their location in the differentiation pathway (12). A recent study using reverse transcriptase polymerase chain reaction (RT-PCR) and immunohistochemical analyses established that mouse mammary stem cells were negative for both ERα and the progesterone receptor (50). An independent study further supported this result by showing that ER-positive cells retained very little stem cell activity (51).

Throughout much of mammary development during embryogenesis, the developing gland consists of a simple duct system, with a primitive mammary bud, the growing terminal end of the duct, with no ERα expression until the 30th week of gestation. These structures are thought to be made of primitive ductal and myoepithelial progenitors. Later in gestation, ERα-positive cells are present in the growing terminal ends of these ducts, followed by a surge of ERα-positive cells in this region after birth (12). In the developing mammary gland, insulin growth factor-1 (IGF-1) is required to form ducts and terminal end buds, indicating its critical role in stimulating the proliferation and differentiation of primitive breast epithelium (12).

ER-positive progenitors arise after 30 weeks of gestation. They respond to estrogen stimulation by proliferating and generating ductal epithelial cells and alveolar cells that form the adult mammary tree. Maintenance of the myoepithelial progenitor pool and the stem cell compartment may rely on paracrine signaling initiated by the ER-positive progenitor cells. Breast carcinogenesis results from mutations affecting stem and progenitor cells. Thus, the phenotype of different subtypes of breast cancer is determined both by the cell of origin and the particular mutations driving carcinogenesis. Thus, ER-negative tumors arise from stem cells or early ER-negative progenitor cells, while ER-positive tumors arise from either ER-negative or ER-positive progenitors (12).

On the basis of these observations, a breast tumor model has been proposed consisting of three types of tumors. The first is an ER-negative tumor termed type 1, which arises from the most primitive ER-negative stem or early progenitor cells. Using classic histopathology and molecular profiling studies, these tumors are poorly differentiated and display a basal phenotype with expression markers of both luminal epithelial and myoepithelial cells. Clinically, these tumors are highly aggressive and are associated with a poor prognosis. Type 2
tumors also arise from early ER-negative stem or early progenitor cells. However, this group includes stem cell mutations, which allow for the differentiation of a specific subset of tumor cells into ER-positive cells, which may range from a low percentage (6–10%) similar to that seen in normal tissue to a high percentage (70–100%). Thus, these tumors are heterogeneous, with intermediate differentiation and prognosis. Type 3 tumors arise from the transformation of ER-positive progenitors. They are predicted to exclusively express luminal markers and have the best prognosis as they contain differentiated cells and respond to antiestrogen therapy (12). Additional breast stem cell markers include CD44, CD24, aldehyde dehydrogenase (ALDH), Oct-4, Bmi-1, and Wnt/β-catenin.

As discussed above, a small population of cancer cells characterized by CD44 expression but low or undetectable levels of CD24 (CD44^+CD24^-/low) have a high tumorigenic capacity when injected into immunodeficient mice (52). In contrast, the rest of the cancer cells, called nontumorigenic breast cancer cells, have little or no tumorigenic ability. Tumors in mice that originate from purified tumorigenic breast cancer cells contain a mixture of both tumorigenic and nontumorigenic breast cancer cells, which confirms that the CD44^+CD24^-/low population shares the capacity for self-renewal with normal stem cells. Gene expression profiling of tumorigenic breast cancer cells was utilized to develop prognostic tools to assess survival in patients with breast cancer and to predict clinical outcomes in breast cancer (34). An invasiveness gene signature of 186 genes was developed to predict the likelihood of a tumor to metastasize. This invasiveness gene signature is associated with the risk of death and metastasis not only in breast cancer but also in lung cancer, prostate cancer, and medulloblastoma. These results suggest the critical relevance of the CD44^+CD24^-/low tumorigenic subclass of breast cancer cells as a pivotal contributor to the association of the invasiveness signature to clinical outcome. In addition, CD44, a cell-surface receptor that mediates cell–cell contacts, promotes the engraftment of leukemic stem cells in the bone marrow and may thus have an additional role as a critical component of the cancer stem cell niche (53).

As a result of the growing interest of identifying new markers for identifying and isolating stem cells, assays based on ALDH activity have been proposed to be a promising alternative in both murine and human mammary stem cells. ALDH is a family of cytosolic enzyme isoforms responsible for oxidizing intracellular aldehydes, including vitamin A, to carboxylic acids. It is highly expressed in hematopoietic progenitors, in intestinal crypt cells, as well as in breast tumor cells (54–56). ALDH1 or (ALDH isoenzyme 1 or ALDH1A1) has been shown to be responsible for the resistance observed in hematopoietic stem cells and breast cancer tumor cells to the alkylating agent, cyclophosphamide (56–58). A novel system has been designed to detect ALDH using a visible light excitable fluorochrome named Aldefluor® (BODIPY-conjugated amino-acetaldehyde, BAAA) metabolized by both murine and human cells (59). Fluorescence-activated cell sorter (FACS) of the normal breast cell population revealed that 6% of this population is ALDH-positive (60). In breast cancer cell
lines, ALDH-positive cells were more frequently observed in basal-like than in luminal cell lines. ALDH-positive cells form mammospheres at a high efficiency and are tumorigenic in NOD/SCID mice (61).

A recent study showed that when acute myeloid leukemic (AML) patient cells were sorted on the basis of Aldefluor staining, three distinct possibilities were observed (62). High levels of aldehyde dehydrogenase (ALDH\textsuperscript{hi}) cells from a first group of patients were identified as normal nonleukemic stem cells, were present in low number, and were able to reconstitute a complete hematopoietic system when transplanted into NOD/SCID mice. However, a second group of patients had ALDH\textsuperscript{hi} cells displaying a higher side scattering pattern than normal stem cells and these cells were also significantly more abundant. These cells correspond to leukemic stem cells. Finally, a third group of patients showed no ALDH\textsuperscript{hi} cells. This study demonstrates that FACS analyses based on the first ALDH isozyme (ALDH1 or ALDHA1) activity can be used to identify and isolate leukemic stem cells (58). In contrast to the toxicity associated with Hoechst 33342, ALDH is relatively mild and allows for cell sorting using viable cell populations, enriched in stem cells. The availability of antibodies derived against ALDH1 allows for the study of their expression in fixed tissue samples using immunohistochemistry (Fig. 4). Immunohistochemistry of 577 human breast carcinoma specimens from two different sources detected 24% and 30% ALDH-positive cells, respectively. In this study, ALDH expression was strongly correlated with tumor size, grade, ER status, human epidermal growth factor receptor 2 (HER2) overexpression, and poor prognosis (60).

OCT-4 is an additional cancer stem cell marker. The OCT-4 oncogene encodes a nuclear protein that belongs to a family of transcription factors containing the POU DNA–binding domain that plays an important role in maintaining the pluripotent state of embryonic stem cells (63). It may prevent expression of genes activated during differentiation, and its expression is downregulated during differentiation. OCT-4 is normally found in the pluripotent stem cells of pregastrulation embryos, including oocytes, early cleavage-stage embryos, and the inner cell mass of the blastocyst, and knockout of OCT-4 causes early lethality in mice because of the absence of an inner cell mass (63). Thus, these activities associated with OCT-4 suggest that it plays a pivotal role in mammalian development and in the self-renewal of embryonic stem cells. The expression of genes such as OCT-4 is potentially correlated with tumorigenesis and may affect some aspects of tumor behavior, such as tumor recurrence or resistance to therapies. OCT-4 is expressed in bladder cancer, as well as in CD44\textsuperscript{+}/CD24\textsuperscript{−} breast cancer–initiating cells (64,65). It may serve as a multifunctional factor involved in major biological processes such as embryonic development, control of differentiation, and stem cell–based carcinogenesis (64).

Bmi-1 is a cancer stem cell marker that belongs to the polycomb group family of proteins, and was first identified as a c-myc cooperating oncogene in murine lymphomagenesis. Bmi-1 is involved in the regulation of diverse biological processes involved in X chromosome inactivation, repression of cellular
Bmi-1 oncoprotein overexpression correlates with axillary lymph node metastases in invasive ductal breast cancer (67). In addition, the Bmi-1 oncogene induces telomerase activity and immortalizes human mammary epithelial cells (68). Bmi-1 regulates self-renewal of normal and malignant human mammary stem cells, stimulates mammosphere formation and colony size in vitro, and increases ductal/alveolar development in humanized NOD/SCID mammary fatpads (20).

The final stem cell marker in this chapter is wingless (Wnt)/β-catenin. The canonical Wnt/β-catenin signaling pathway modulates the delicate balance between stemness, differentiation, and tumorigenesis in several adult stem cell

Figure 4 ALDH1 immunostaining. Immunohistochemistry using anti-ALDH1 on tissue microarray samples from two patients (A, B) with infiltrating ductal carcinoma of the breast, both magnified 200×, with a scale bar of 100 μm. Abbreviation: ALDH1, aldehyde dehydrogenase isoenzyme 1.
niches such as the hair follicles in the skin, the mammary gland, hematopoiesis, and the intestinal crypt (8,69). Differences in the levels of Wnt signaling activity reflect tumor heterogeneity and are likely to account for distinct cellular activities such as proliferation and epithelial-mesenchymal transitions, which prompt tumor growth and malignant behavior, respectively. This pivotal pathway is highly conserved in evolution and is known to regulate cell-fate decisions, cell proliferation, morphology, migration, apoptosis, differentiation, and stem cell self-renewal. Recent evidence which suggests that Wnt signaling plays a role in human breast cancer include the detection of elevated levels of nuclear and/or cytoplasmic β-catenin using immunohistochemistry, and the overexpression or downregulation of specific Wnt proteins. The Wnt pathway has also been implicated in normal stem cell self-renewal in vivo, and there is evidence that dysregulation of this pathway in the mammary gland and other organs may play a key role in carcinogenesis (70).

The Wnt family of secreted proteins includes the well-characterized canonical Wnt signaling pathway, in which Wnt ligands signal through the stabilization of β-catenin. In this pathway, Wnt proteins bind to a family of frizzled receptors in a complex with the low-density lipoprotein receptor–related proteins 5 and 6 (LRP5/6) coreceptors to activate Dishevelled (Dsh). Subsequently, Dsh inhibits the activity of the β-catenin destruction complex [adenomatous polyposis coli (APC), axin, and glycogen synthase kinase-3β (GSK-3β)], which phosphorylates β-catenin in the absence of the ligands. As a result, β-catenin is stabilized and translocated to the nucleus, where it recruits transactivators to high mobility group (HMG)–box DNA-binding proteins of the lymphoid-enhancer factor/T-cell factor (LEF/TCF) family. In the absence of Wnt signaling, β-catenin remains in the cytoplasm, where it forms the β-catenin destruction complex. GSK-3β phosphorylates β-catenin, which targets the protein for ubiquitin-mediated degradation. When the Wnt pathway is activated, GSK-3β is inhibited, blocking β-catenin phosphorylation and its subsequent degradation (71). In addition, several β-catenin-independent Wnt signaling pathways, known as noncanonical, have been shown to be crucial for different aspects of vertebrate embryo development (71).

**MOLECULAR BREAST CANCER STEM CELL TARGETS**

Molecular signatures of stem cells are currently being identified. The mutations that accumulate in cancer stem cells could inappropriately activate Bmi-1, Wnt/β-catenin, Notch, sonic hedgehog (SHH), and PTEN pathways that are involved in the regulation of self-renewal of normal stem cells and could potentially lead to cancer (Figs. 5 and 6) (26,72–74). These molecular markers can be exploited to purify cancer stem cells, which can then be characterized in the most intimate detail, by the modern tools of gene expression profiling and informatics. The failure to shut off the self-renewal pathways could confer self-renewal on cells and micro-environments capable of evading homeostatic regulation, leading to tumorigenesis.
The intimate interrelationship of multiple pathways critical to tumorigenesis and to normal stem cell self-renewal raises the specter that cancer therapies targeting such pathways might ablate resident stem cells (75). Drugs that selectively target the self-renewal pathways of tumors will be attractive candidates for cancer therapies if they can effectively spare the normal bystander stem cells that use self-renewal as part of their regular functional repertoire. Alternatively,

Figure 5  Stem cell self-renewal pathways include the HH, Notch, Bmi-1, Wnt, and PTEN pathways. Dysregulation of the self-renewal pathways of NSCs and progenitor cells may lead to CSCs and tumor formation. NSCs are ER negative and they differentiate into ER-positive progenitor cells, which further differentiate into myoepithelial, alveolar epithelial, and ductal epithelial cells. CSCs in their niche may lose asymmetric division, leading to the clonal expansion of cancerous cells and tumor formation. Abbreviations: NSCs, normal stem cells; CSCs, cancer stem cells; ER, estrogen receptor.

The intimate interrelationship of multiple pathways critical to tumorigenesis and to normal stem cell self-renewal raises the specter that cancer therapies targeting such pathways might ablate resident stem cells (75). Drugs that selectively target the self-renewal pathways of tumors will be attractive candidates for cancer therapies if they can effectively spare the normal bystander stem cells that use self-renewal as part of their regular functional repertoire. Alternatively,
Figure 6  A model of the interrelationship of stem cell self-renewal pathways including a conceptual link between sporadic and hereditary breast cancers. BRCA1 may inhibit stem cell self-renewal, while HER2 overexpression may increase the stem cell pool through regulation of self-renewal. In this model, there is bidirectional cross talk between the Notch and HH pathways with subsequent regulation of the polycomb gene Bmi-1 and the Wnt/β-catenin pathway. SHH, IHH, and DHH are HH ligands, while Gli1 and 2 are two HH transcription factors that are positive mediator of HH signaling. Cyclopamine inhibits the HH pathway. The DSL peptide activates Notch signaling, while the GSI blocks the intramembrane cleavage of Notch required for signaling. PTEN affects Wnt signaling through GSK-3β. Exisulind, bromoindirubin-3'-oxime, and imatinib inhibit the Wnt/β-catenin pathway. Abbreviations: BRCA, breast cancer; HER, human epidermal growth factor receptor 2; HH, hedgehog; SHH, sonic hedgehog; IHH, Indian hedgehog; DHH, desert hedgehog; PTEN, phosphatase and tensin homolog deleted on chromosome ten; GSI, gamma secretase inhibitor.
differentiating therapies may also be effective if they can efficiently shift the balance of stem cell divisions toward differentiation rather than self-renewal and stem cell expansion. In addition to Bmi-1, and Wnt/β-catenin discussed above, Notch, SHH pathways, and PTEN will be reviewed below.

The Notch signaling pathway plays an important role in regulating proliferation, survival, and differentiation of many cell types, including the regulation of self-renewal of adult stem cells and differentiation of progenitor cells along a particular lineage. These outcomes include increased survival or death, proliferation or growth arrest, and commitment to or blockage of differentiation. There are four mammalian Notch homologues, Notch 1 to 4, which interact with the DSL ligands: Delta, Delta-like, Jagged1, and Jagged2 in vertebrates (71).

Notch 1 and Notch 4 homologues are involved in the normal development of the mammary gland, and mouse mammary tumors are associated with their mutated forms. Notch signaling is critical to normal human mammary development where it acts on both stem cells and progenitor cells, affecting self-renewal and lineage-specific differentiation. Notch signaling may also contribute to mammary carcinogenesis by dysregulating the self-renewal pathways of normal mammary stem cells. There is bidirectional cross talk between the Notch and HH pathways with subsequent regulation of the polycomb gene Bmi-1, as well as the Wnt signaling network (71). The DSL ligands activate Notch signaling, while the gamma secretase inhibitor (GSI) blocks the intramembrane cleavage of Notch required for signaling (71). The abnormal expression of Notch receptors in epithelial metaplastic lesions and neoplastic lesions suggests that Notch may act as a protooncogene (26). In spite of the simplicity of the Notch signaling pathway, it yields varied outcomes in different contexts (76). These different outcomes are mediated through a novel signaling pathway in which Notch receptors on the cell surface are processed to give rise to a nuclear transcriptional activation complex. Activated Notch modulates stem and progenitor cell renewal through the upregulation of the cyclin-dependent kinase inhibitors p21cip1 and p27kip1 (77). Notch signaling is also required for the survival of niche cells (78). Notch receptor activation promotes the survival of neural stem cells by inducing the expression of the specific target genes hairy and enhancer of split 3 (Hes3) and SHH. The underlying mechanism occurs through the rapid activation of cytoplasmic signals, including the serine/threonine kinase AKT, the transcription factor STAT3, and mammalian targets of rapamycin (mTOR) molecular targets of rapamycin (79).

The HH signaling pathway has important roles in embryonic development and tumorigenesis (80). In a number of tumors, including prostatic and pancreatic cancers, HH signaling is abnormally upregulated, and effective signaling inhibition can repress tumor growth and induce apoptosis (81–83). The core components of the HH pathway are evolutionarily conserved from insects to mammals. They include the secreted ligand HH, the membrane receptor Patched (PTCH), the membrane signal transducer Smoothened (SMO), and the transcription factor Ci/glioma (GLI). However, while in Drosophila, there is only
one secreted ligand (HH), one receptor (Ptc), and one transcription factor (Ci), there are three secreted ligands, sonic HH (SHH)/desert HH (DHH)/Indian HH (IHH), two receptors (PTCH1/2), and three transcription factors (GLI1/2/3) in mammals (81). The HH signaling components PTCH1 and GLI1/2 are highly expressed in normal mammary stem and progenitor cells, and activated HH signaling is mediated by Bmi-1 (30). Addition of Hh ligands increases the expression of the transcription factors Gli1 and Gli2, which are positive mediators of Hh signaling (20). SHH can potentiate tumor cell proliferation. Components of the SHH–GLI pathway are expressed in adult human prostate cancer, often with enhanced levels in tumors versus normal prostatic epithelia. Cyclopamine and anti-SHH antibodies inhibit the proliferation of GLI1–PSA primary prostate tumor cultures (83). The SHH signaling pathway also regulates adult neuronal self-renewal, and may play a role in brain tumor biology (4).

The final pathway discussed herein is PTEN, which is a phosphatase that inhibits cellular proliferation by negative regulation of signaling through the phosphatidylinositol-3-OH kinase (PI3K) pathway. The activation of PI3K leads to phosphorylation and activation of the Akt protein, which in turn can potentially phosphorylate a multitude of proteins (84). The PTEN–Akt pathway regulates stem cell activation by helping control nuclear localization of the Wnt pathway effector β-catenin. PTEN inhibits, whereas Akt enhances, the nuclear localization of β-catenin. Akt is thought to coordinate with Wnt signaling to assist in the activation of β-catenin in intestinal stem cells, either through phosphorylation of GSK-3β or by phosphorylation of β-catenin itself. The phosphorylation by Akt inactivates GSK-3β, helping the Wnt signal to permit β-catenin to escape proteasome-mediated degradation and facilitating its nuclear translocation and activity in transcriptional regulation (85).

In addition to other downstream effectors, the highly branched PI3K pathway activates the mammalian target of rapamycin (mTOR). Molecules such as p21Cip1/Waf1 are found downstream to PTEN, and they influence stem cell proliferation, without necessarily implicating self-renewal. FOXO regulatory proteins may have the effects of PTEN on p21Cip1/Waf1 by influencing the expression of a range of crucial genes for stem cell function, including cell division, and tolerance of oxidative stress. Thus, the perturbation of PTEN may thus result in nonrenewing cell divisions or death of stem cells (84). PTEN has also been reported to be a guardian of genome integrity and of the maintenance of chromosomal stability through the physical interaction with centromeres and control of DNA repair. PTEN acts on chromatin and regulates the expression of Rad51, which reduces the incidence of spontaneous double strand breaks (86).

PTEN is commonly deleted or otherwise inactivated in diverse cancers, including hematopoietic malignancies, glioblastomas, endometrial carcinomas, breast carcinomas, and prostate carcinomas (86). PTEN deletion or HER2 overexpression may modulate stem cell function leading to increased invasion and tumorigenicity (87). PTEN deletions are implicated in the generation of transplantable leukemia-initiating cells, and the depletion of normal hematopoietic
stem cells (88). PTEN normally maintains hematopoietic stem cells in a quiescent state, and in the absence of PTEN, hematopoietic stem cells are driven into the cell cycle, eventually leading to depletion of hematopoietic stem cell reserves. Thus, PTEN activity identifies a mechanistic difference between the maintenance of normal stem cells and cancer stem cells. These effects were mostly mediated by mTOR as they are inhibited by rapamycin, which depletes leukemia-initiating cells but also restores normal hematopoietic stem cell function (88). This mechanistic distinction shall enable the design of therapies that target the PTEN pathway to combat leukemia with the goal of killing leukemic stem cells, without damaging the normal stem cell pool (75). In addition to leukemia, this pathway may be an attractive therapeutic target for treating a broad spectrum of tumors carrying inactivated PTEN because mTOR activity can be suppressed by a number of compounds including rapamycin (75). Genes that promote stem cell quiescence might be particularly amenable to such targeted strategies because although normal stem cells thrive in a quiescent world, forcing cancer stem cells to adopt quiescence could lead to their eradication (75).

TARGETS OF THE STEM CELL NICHE

While cancer stem cells may indeed be lethal seeds in themselves, they may whither and die unless they are supported by environmental factors, which perpetuate their survival and self-renewal. The stem cell niche provides this microenvironment for cancer stem cells, and the combination of lethal seeds in lethal soil is likely to lead to the most aggressive forms of breast cancer, with the most fatal course. The stem cell niche determines cell fate decisions based on surrounding cells, extracellular matrix components, and secreted local and systemic factors (11). The 1889 proposal by Stephen Paget that metastasis depends on cross talk between selected cancer cells (the “seeds”) and specific organ microenvironments (the “soil”) still holds forth today, and may also apply to the interrelationship between cancer stem cells and their niche (89–91). Cancer metastasis requires the seeding, homing, and successful colonization of specialized cancer stem cells at distant sites (92). In this context, it is not surprising that the inhibition of the homing factor CXCR4 has been shown to effectively prevent both primary tumor formation as well as metastasis in animal models (92). Stem cells of various tissues exist within protective niches that are composed of differentiated cells which provide direct cell contacts and secreted factors that maintain stem cells primarily in a quiescent state by providing cues for the inhibition of proliferation (8,44). The mutations that distinguish cancer stem cells from normal stem cells may enable them to escape niche control. Alternatively, the niche provides proliferation or differentiation-inducing signals at times when high numbers of progenitors are needed and can quickly give rise to all committed cell lineages (8). As a result, the uncontrolled proliferation of stem cells and tumorigenesis may be a consequence of the dysregulation of extrinsic factors within the niche. In this case, the stem cell niche might represent
targets for cancer therapy (44). Given that postnatal estrogen exposure is related to breast cancer risk, this effect could also apply in utero, where the developing fetus is exposed to estrogen levels up to 10-fold higher than at any other time in the individual’s lifetime. Under these circumstances, estrogens or other hormones may create the fertile soil that is responsible for an increase in the responsiveness of estrogen-responsive tissue to undergo oncogenesis later in life (93).

Epigenetic modulation of gene expression is highly abnormal in cancers, which often show aberrant promoter CpG island hypermethylation and transcriptional silencing of tumor suppressor genes and prodifferentiation factors. The aberrant epigenetic landscape of the cancer cell is also characterized by a massive genomic hypomethylation, an altered histone code for critical genes, and a global loss of monoacetylated and trimethylated histone H4 (94). In embryonic stem cells, these genes are maintained in a state that is ready for transcription mediated by a bivalent promoter chromatin pattern consisting of the repressor, histone H3 methylated at Lys27 (H3K27) by polycomb group proteins, plus the activator, methylated H3K4. However, in contrast to embryonic stem cells, embryonic carcinoma cells add two key repressors, dimethylated H3K9 and trimethylated H3K9, which are both associated with DNA hypermethylation in adult cancers (14). Polycomb group proteins reversibly repress genes required for differentiation in embryonic stem cells. Stem cell polycomb group targets are up to 12-fold more likely to have cancer-specific promoter DNA hypermethylation than nontargets, supporting a stem cell origin of cancer where reversible gene repression is replaced by permanent silencing, locking the cell into a perpetual state of self-renewal, thereby predisposing to subsequent malignant transformation (95). Thus, cell chromatin patterns and transient silencing of these important regulatory genes in stem or progenitor cells may leave these genes vulnerable to aberrant DNA hypermethylation and heritable gene silencing during tumor initiation and progression.

Vascular niches have been described in the case of neural stem cells, which were found to be concentrated in regions that are rich in blood vessels (18). These vessels shelter neural stem cells from apoptotic stimuli and maintain a proper balance between self-renewal and differentiation. The endothelial cells that line blood vessels are key components of the vascular niche as they secrete factors that promote stem cell survival and self-renewal. As the cotransplantation of tumor cells with endothelial cells leads to more rapid tumor formation, the vascular niche may contribute to tumor initiation and development (44). In an established tumor, cancer stem cells that find a vascular niche may continue to self-renew, while the remaining cells may differentiate to form the bulk of the tumor, but not to its long-term maintenance (18). Thus, the identification of endothelial cell signals, which regulate the self-renewal of cancer stem cells, may also contribute to our understanding of the interrelationship of cancer stem cells and their niche. There is some evidence that there is a bidirectional relationship between cancer stem cells and their niche. For example, gliomas secrete elevated levels of vascular endothelial growth factor (VEGF), which increases
endothelial cell migration and tube formation (43). In addition, VEGF promotes neural stem cell survival (96). Thus, the effectiveness of antiangiogenic agents in cancer therapy may relate to the prevention of new blood vessels that inhibit tumor growth. In addition, antiangiogenic therapy may disrupt a vascular niche, which is necessary for cancer stem cell renewal (18).

In the case of hematopoietic stem cell niches, studies have focused on the endosteum, the inner surface of the bone that interfaces with the bone marrow. Regulation of hematopoietic stem cell function is governed by factors secreted by osteoblasts, osteoclasts, and stromal fibroblasts within the endosteum (8). Imaging data suggests that hematopoietic stem cells reside near the endosteum, and that this localization and maintenance is also usually adjacent to CXCL12-secreting reticular cells, irrespective of whether the hematopoietic stem cells are in vascular or endosteal locations (97,98). The stimulation via parathyroid hormone–related protein receptors (PPR) or an increase in number of osteoblastic cells in the periosteum with high levels of the Notch ligand Jagged1, enlarged the stem cell niche with a resultant increase in the number of hematopoietic stem cells with evidence of Notch1 activation in vivo (99). The activation of Notch1 has been shown to result in enhanced self-renewal of hematopoietic stem cells possibly by inducing self-renewal genes including Hes1 (8). In addition, the bone morphogenetic protein signal has an essential role in inducing haematopoietic tissue during embryogenesis within the hematopoietic stem cell niche (100). Mesenchymal stem cells have been shown to improve engraftment of hematopoietic stem cells, and suppress graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. In addition, when tumor cells are injected into NOD/SCID mice in conjunction with mesenchymal stem cells, their growth is much faster as compared to the group receiving only tumor cells. To explain this observation, it was suggested that mesenchymal stem cells have the ability to form a cancer stem cell niche in which tumor cells can preserve their proliferative potential to sustain the malignant process (101).

Activated, irradiated, mammary fibroblasts are critical to the microenvironment of mammary stem cells as they are required for the successful repopulation of cleared mammary fat pads in immunocompromised mice (102). These epithelial-stromal interactions are also important in malignant transformation and likely to be important with cancer stem cells (11). Fibroblasts secrete matrix metalloproteinases and protease inhibitors and mobilize and activate growth factors such as tumor growth factor-β (TGF-β), which has been shown to promote tumor progression and invasion (103). Carcinoma-associated fibroblasts promote angiogenesis, as well as tumor growth, by secreting stromal cell-derived factor-1 (SDF-1), which interacts with the chemokine receptor CXCR4 on tumor cells and endothelial cells (104). As radiographic breast density correlates with increased breast cancer risk, as well as increased fibroblasts and collagen deposition, this modified niche may promote the transformation of initiated stem cells into cancer stem cells (11,105).
FUTURE CHALLENGES: INDIVIDUALIZED THERAPY BASED ON BREAST CANCER STEM CELLS AND THEIR NICHE

In the last few years, there has been a significant evolution in breast cancer care, if not a revolution (106). Whereas, we previously focused on tumor size, nodal status, and certain prognostic features based on tumor pathology, the paradigm has now shifted to include the underlying molecular pathways that drive cancerogenesis, as well as the cellular targets of breast cancer, including cancer stem cells and their niche. The intricate circuitry of pathways that define the hallmarks of cancer and the corresponding complex molecular interactions are being translated into therapeutic benefit (107,108). Individualized drivers are gradually being factored into patient assessments in order to astutely and accurately determine the clinical course of breast cancer patients on the basis of their unique and complex cellular and molecular breast cancer portraits.

Breast cancer stem cell markers and the critical molecules in the pathways that drive their self-renewal are necessary to complete the molecular portrait of breast cancer. Our failure to eradicate most cancers may be as fundamental as a misidentification of the target cancer stem cell (Fig. 3) (6). The fact that stem cells rarely divide and have unique cellular properties, coupled with the observation that they may have high levels of drug transporters to pump chemotherapy agents out of the cell, has led many to believe that traditional chemo- or radiation therapies are insufficient to clear these tumor-initiating cells from the body (92). Clinical observations indicate that breast cancer responses to therapy do not necessarily correlate with patient survival and are incomplete clinical endpoints. Instead, the effectiveness of these treatments should be evaluated by decreased cancer recurrence and metastases as measures of therapeutic efficacy (92). As the molecular pathways that govern normal and cancer stem cells overlap, the targeting of these pathways could have harmful effects on the homeostatic function of normal stem cells. Targeted therapies that specifically eliminate cancer stem cells could further revolutionize the ways by which we treat cancer.

As the proliferation of cancer stem cells is important for the growth of the primary tumor, as well as progression to metastatic disease, inhibition of the self-renewal capacity of these cells is an obvious target for their elimination. Therapies such as cyclopamine (targeting the HH signaling) and exisulind, bromoindirubin-3′-oxime, and imatinib (targeting the Wnt/β-catenin pathways) were designed to inhibit self-renewal pathways critical to cancer stem cells, and have achieved varying levels of success (92). Another strategy to eliminate self-renewal is to force cancer stem cells to differentiate (Fig. 6).

As cancer stem cells are currently being identified in a broad spectrum of cancers (44,109–112), significant advances in their fundamental properties are forthcoming, and will serve as a scaffold to expedite our understanding and identification of cancer stem cells within specific types of cancers, including breast cancer. The existence of overlapping cancer stem cell pathways within different cancer types will also stimulate the discovery and development of new drugs to
target these pathways, as they are likely to be effective and marketable in more than merely a small subset of cancer patients.

The addition of molecular markers in our patient assessments acknowledges the pivotal role that the corresponding molecular pathways play in fueling tumorigenesis. This fine-tuning of patient prognostic factors will ensure that the oncogenic drivers of each individual patient’s tumor pathophysiology are both identified and tailored into the definition and optimization of their specific therapeutic strategy, which will translate into their best possible clinical outcome and overall survival. As a result, it is critical that our therapeutic armamentarium target the specific fingerprint of cellular and molecular pathways defined for an individual patient. Therapeutic resistance via quiescence and the sequestration of cancer stem cells into sanctuaries that are impermeable to most therapeutic agents must be taken into consideration in the design of individualized breast cancer therapies. Therapeutic resistance may also result from an inadequate depletion of key cellular and molecular elements within the cancer stem cell niche. The elucidation of the secreted factors that set up the premetastasis and metastatic niche could provide both diagnostic as well as therapeutic benefits.

Drugs such as Bevacizumab (Avastin) that reduce intratumoral pressure associated with leaky tumoral vasculature and thereby increase the introduction of intravenous agents within the tumor bed may have the critical synergy necessary in combinations of targeted agents to permeate cancer stem cell sanctuaries and cancer stem cell niches. In addition, they may target both the vascular niche and the associated cancer stem cells. Thus, it is not surprising that clinical trials of Bevacizumab combined with the chemotherapeutic drug CPT-11 suggest that this combination is one of the most effective treatments of glioblastoma. Other agents that stimulate cycling of quiescent cancer stem cells may also be beneficial in combined therapies, which include agents that eradicate cycling cells. In addition, individualized therapy must also have the long-term objectives of predicting and preparing to eradicate the molecular pathways that may circumvent those pathways that drove the initial breast cancer pathophysiology, and were targeted by first-line targeted therapies. This strategy also underlines the critical objective of eradicating every single cancer stem cell and its niche in newly diagnosed breast cancer patients, thereby striving to eliminate all possibilities that these lethal seeds in lethal soil may ignite tumor recurrence in the future.

Therefore, the genesis of individualized breast cancer therapy including the eradication of cancer stem cells has become a necessity to achieve optimal breast cancer patient care. It is imperative to understand the biology of cancer stem cells and their niche to optimize our cancer treatment strategies to prevent the longevity of cancer stem cells from interfering with our own. The associated costs of its future reality precipitate a call to rethink our healthcare strategies with a focus on cancer prevention and therapeutic schedules which deliver targeted agents when they are both necessary and most effective in our quest to optimally treat an individual breast cancer patient with the ultimate goal.
Breast Cancer Stem Cells and Their Niche: Lethal Seeds in Lethal Soil

of a stable, long-term remission. In this setting, the identification and targeting of breast cancer stem cells and their niche hold great promise in the development of new therapeutic translational strategies using state-of-the-art individualized therapies for breast cancer patients.

ACKNOWLEDGMENTS

Rhyna Salinas (CHUM) is gratefully acknowledged for her expert assistance in the preparation of this chapter. The contributions made by Drs. Francine Jolicoeur, Isabel Garcia de la Fuente, and Louis Gaboury (CHUM) to Figure 4 during their collaboration with D.B. are greatly appreciated. Drs. Connie Eaves, Gabriela Dontu, Suling Liu, Christophe Ginestier, Hasan Korkaya, and Julie Dutcher are thankfully acknowledged for insightful and inspirational discussions. Dr. Connie Eaves is also gratefully recognized for her invaluable role as D.B.’s first stem cell research mentor and role model, who initially planted the seed that she continues to nourish. D.B. is a scholar of Fonds de la recherche en santé du Québec (FRSQ) and principal investigator of the National Cancer Institute of Canada’s Clinical Trials Group (NCIC CTG). This chapter was also made possible through a Canadian Institutes for Health Research (CIHR) operating grant to D.B..

REFERENCES

Breast Cancer Stem Cells and Their Niche: Lethal Seeds in Lethal Soil

Molecular Imaging in Individualized Cancer Management

David M. Schuster
Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, Emory University, Atlanta, Georgia, U.S.A.

Diego R. Martin
Department of Radiology, Emory University, Atlanta, Georgia, U.S.A.

INTRODUCTION

X-rays were discovered by Karl Roentgen in 1895, and with that discovery came the ability to see structures within the living human body without having to directly sample the tissue. Imaging, particularly cross-sectional techniques, has revolutionized the practice of medicine. These techniques employ ionizing radiation, such as computed tomography (CT) scanning, single-photon emission computed tomography (SPECT), and positron emission tomography (PET), as well as non-ionizing energy, such as ultrasound and magnetic resonance imaging (MRI). Some of the most exciting and practical major advances are unfolding within the MRI and PET technologies, specifically with regard to molecular imaging.

Molecular imaging is defined as the in vivo characterization and measurement of biological processes at cellular and molecular levels. By definition, molecular imaging is not restricted to any one tool; it may involve optical, ultrasound, gamma ray, positron, or magnetic imaging (1). Since the 1960s, bone scanning has played a major role in the management of breast cancer. However, in the past decade, the role of radionuclide molecular imaging has expanded
significantly in the clinical management of breast cancer because of PET, mammography, and sentinel lymph node techniques. Molecular imaging is also instrumental in drug development, gene therapy, and basic science research of breast cancer. Mammography and ultrasound have been and continue to be very important tools in the diagnosis and management of breast cancer, but other modalities, such as MRI, have enhanced our ability to detect occult disease, especially in those with high genetic risk, and to follow breast carcinoma during and after therapy (2,3).

In this chapter, we will examine MRI and PET with regard to molecular imaging as applied to individualized breast cancer management; this chapter is not intended to be a comprehensive review of molecular imaging methods. We have specifically chosen MRI and PET because we believe the integration of both technologies will lead to further progress in our ability to interrogate disease processes in a manner analogous to the classical pathologist, but in a noninvasive fashion. This ability will continue to expand as technology progresses. But in essence, the future is already here.

RADIONUCLIDE TECHNIQUES: CURRENT CONCEPTS AND APPLICATIONS

The advantage of radionuclide techniques over other imaging approaches is their ability to label any chemical species with an isotope of choice and to detect that radiotracer with high sensitivity. When certain radionuclides decay, gamma rays (photons) are given off, which are detected with special crystal detectors. The gamma emissions of SPECT or PET radiotracers are energetic enough to travel through tissues and thus enable imaging of the whole body. Carbon, nitrogen, hydrogen, and oxygen have positron-emitting isotopes and are ideally suited for designing probes to image natural biochemical and molecular process in vivo. Probes can either home in on targets on cell surfaces, such as antigens and receptors, or intracellular structures such as internalized receptors.

Single-photon imaging is the basis of simple techniques such as planar bone scanning or more sophisticated cross-sectional approaches such as SPECT. High-energy dual-photon imaging is utilized for PET, permitting better resolution and sensitivity than single-photon techniques. Yet, each method has its place. For example, sentinel lymph node procedures, employing single-photon radiotracers, are rapidly becoming the standard of care in breast cancer lymph node staging (4). For all practical purposes modern PET scanning is now accomplished as part of a PET-CT device, which fuses the functional imaging of PET with the anatomic detail of CT.

In 1984, Beaney and coworkers reported (5) on the use of [15O]H2O PET to study blood flow, the oxygen extraction ratio, and oxygen utilization in breast cancer. The earliest studies on the use of 18F-fluorodeoxyglucose (FDG) in the diagnosis of breast cancer were conducted in 1989 (6,7). Figure 1 is an example of 18F-FDG PET uptake in a primary breast mass and axillary lymph nodes.
Fluorodeoxyglucose is actively transported into the cell, where it is irre-
versibly phosphorylated by hexokinase, trapping it within the cell, where it does
not proceed further down the glycolytic pathway. There is a correlation between
FDG uptake and glucose transporter-1 (Glut-1) expression, the mitotic activity
index, percent necrosis (likely reflecting the proliferation rate), tumor cells/
volume, hexokinase-1 expression, the number of lymphocytes (but not macro-
phages), microvessel density, and possibly the expression of the p53 gene (8–12).

PET is currently of limited value in routine screening for breast cancer (13)
and cannot take the place of fine sectioning and immunohistochemical lymph
node evaluation, but it has an important role to play in a select group of prob-
lematic diagnostic situations and in staging patients with high-risk disease (14–17).
PET can provide important information by detecting locoregional lymph node
involvement as well as distant metastases. Because of its high positive predictive
value, FDG PET may obviate a sentinel lymph node procedure and axillary
dissection.

The use of neoadjuvant chemotherapy has increased the rate of breast
conservation surgery (18). Clinical response does not necessarily correlate with
pathologic response. FDG PET has been shown to aid in monitoring response to
chemotherapy and has prognostic benefit (19). Nonresponders and those
developing progressive disease or distant metastases can be identified earlier,
and this information may prove useful in changing therapies and avoiding side
effects of chemotherapy that is not effective. Baseline PET combined with PET
after the first course of chemotherapy is highly accurate in this regard (16). A
meta-analysis of the FDG PET literature (20) reports an overall 81% sensitivity,
96% specificity, and 92% accuracy for monitoring response to therapy. Elevated
FDG uptake was found to be an independent predictor of worse prognosis and
decreased relapse-free survival in breast cancer (10). Figure 2 illustrates a dramatic
response of metastatic breast cancer to fulvestrant (Faslodex) therapy.

FDG PET has been shown to locate unsuspected metastases with high
sensitivity and specificity (21,22). Detection of early recurrence may have

Figure 1  (A) Axial CT, (B) attenuation corrected PET, and (C) fused image in a 50-year-old
female with a 4-cm primary right breast cancer (arrow) and malignant axillary lymph
nodes (arrowhead). Abbreviations: CT, computed tomography; PET, positron emission
tomography.
important survival benefits, prompting the use of new therapies as well as curative or palliative surgery. FDG PET is considered of great efficacy in the evaluation of patients with suspected recurrent breast cancer, especially to differentiate true recurrence from postsurgical and radiation sequelae. The ability of PET to detect suspected clinical recurrence and in patients with elevated tumor markers has been demonstrated in a number of studies (23–32).

MRI: CURRENT CONCEPTS AND APPLICATIONS

In 1973, Paul Lauterbur discovered that introduction of magnetic gradients superimposed over a magnetic field makes it possible to create two-dimensional images of structures, with the first experiments showing the ability to differentiate between tubes filled with regular or heavy water (33). No other imaging method can achieve this differentiation. This work was the foundation for a revolutionary tissue-imaging technique that generates images on the basis of molecular differences between or within tissues.

To understand the added value of MRI to clinical sciences, it is useful to compare it with CT. CT can produce images of high spatial resolution, which are interpreted on the basis of anatomy. A distinct feature of MRI is that anatomically correct images are produced, but these images are based on molecular events that produce much higher soft tissue contrast between tissues and between
abnormal and normal tissues. It has been shown for most clinically important applications, including the brain and the solid organs of the body, that soft tissue contrast resolution is more important than spatial resolution. Even on contrast-enhanced CT, the ability to generate soft tissue contrast resolution is less than that of MRI, and this difference may contribute to the relative advantage that MRI provides for lesion detection (Fig. 3).

Unlike CT, which relies on the degree of X-ray absorption to differentiate tissues, MRI contrast differences are based on molecular interactions between protons located in water molecules within and between cells, neighboring molecules that influence mobility, such as charged glycoproteins and membranes, and species that can alter the magnetic field at the molecular level, such as iron and lipid. There are essentially an infinite variety of formulations of MRI techniques that can exploit these interactions to generate different contrast patterns. Similar to the concept of a pathologist using different histochemical stains to visualize molecular species, the expert MR imager will utilize an optimized set of MRI sequences to visualize the characteristics that may be unique to specific pathologies. Though the initial applications of MRI had been for imaging diseases of the brain, MRI is now able to image faster and with concomitant improvement in image quality. This has been critical for improving the capacity of MRI for rapidly imaging the chest, abdomen, and pelvis, where motion from respiration or cardiac contractions may deteriorate images (34,35).

Breast MRI (Fig. 4) is a recent additional imaging method for the evaluation of primary breast carcinoma and is useful for early detection and staging (36–40). MRI has been found to provide maximal sensitivity, as compared to traditional mammography and ultrasound, but with comparatively lower specificity. The fundamental components to the current state of the art of breast MRI is to use the combination of time-resolved gadolinium uptake analysis in combination with the evaluation of the contours of a mass using high-contrast and
Figure 4  (A) Right breast carcinoma shown on sagittal plane dynamically enhanced gadolinium MRI at the indicated time intervals. Subsequently, (B) a region of interest is applied within the tumor, and a time-intensity curve is generated. (C) A lymph node showing a similar enhancement pattern (arrow) is shown on a more lateral slice through the right axilla on a representative 3-min sagittal image. (D) The enhancing primary breast carcinoma and the metastatic lymph node in the right axilla can be seen on an axial MRI image of both breasts (arrows), which correlates with findings on (E) PET-CT (arrows), confirming the presence of metastasis in the node. Note that the position of the breast is different on the MRI compared with the PET-CT due to prone MRI versus supine PET positioning. Abbreviations: MRI, magnetic resonance imaging; PET, positron emission tomography; CT, computed tomography.

high-resolution imaging. MRI techniques have benefited from improvements in surface coils specifically dedicated to imaging both breasts simultaneously.

The most important sequence used for routine breast MRI is the three-dimensional gradient echo (3D GRE) technique. Uniform and complete suppression of the breast lipid signal has been a key area of focus for improvement of these sequences. With time-resolved imaging after an MRI contrast injection (Fig. 4), malignant focal breast carcinomas typically show enhancement that
remains constant over several minutes or may diminish. Benign tumors such as fibromas may show differential contrast uptake, with progressive increase in contrast accumulation within the benign mass. However, the specificity for these differentiating features is relatively moderate. A more recent approach has incorporated lessons learned previously from conventional X-ray mammography; the margins of the tumor may show patterns specific to malignancy. Data are developing to support that the analysis of tumor margin morphology will improve MRI specificity for the diagnosis of breast carcinoma and separate this condition from benign tumors. Here, further development of newer MRI strategies, such as perfusion studies or spectroscopy, or a possible combination with metabolic imaging provided by PET, may help improve this apparent current limitation.

Lymphatic spread may result in local lymph node involvement. Findings on MRI may be relatively specific when the lymph nodes are found to be over 1 cm in diameter and show contrast enhancement that is similar in time course and intensity to those seen in the primary breast mass. However, for smaller lymph nodes the specificity may be significantly lower. Combined imaging with FDG-PET and MRI may help aid the specificity of the MRI study and the identification of the most concerning lymph nodes (Fig. 4).

Currently accepted applications for MRI include screening for high-risk women, evaluation of dense fibroglandular breast tissue, and surveillance after surgical alteration or breast implant placement (36–40). Generalized use of MRI for evaluation of all women in the screening age group has not been widely accepted.

With regard to hematogenous tumor spread, the liver is the organ most commonly involved in metastatic diseases, usually from breast, colon, pancreas, kidney, and carcinoid cancers, or melanoma. However, most liver tumors are benign, and benign liver lesions are common, including cysts, bile duct hamartomas, hemangiomas, adenomas, and focal nodular hyperplasia. The specificity of a noninvasive test is the litmus test for utility in tumor evaluation. We have shown that although MRI is slightly more sensitive than CT for tumor detection, the most important distinction is that MRI is highly specific, approaching 98%, while CT is significantly less specific, approaching 80% (41–44). The ability to image rapidly and include the evaluation of the entire body is undergoing development. The potential for assessing metastatic burden will be a primary potential role for this capability.

FUTURE DEVELOPMENTS

The resolution of PET scanners is currently limited in comparison with anatomic imaging, but continued technological improvements are expected (45,46). Sensitivity for initial diagnosis will likely improve as device resolution advances and as special focused field PET-mammography imaging systems are developed (47–49). Small field-of-view cameras not only offer better sentinel
lymph node and primary tumor detection because of increased spatial resolution but can also be mounted on maneuverable arms and used in an intraoperative setting, helping the surgeon realize a tumor-free margin (50).

MR spectroscopic analysis has the potential to elucidate pathophysiological processes and may become useful for the assessment of primary or metastatic neoplasms, or the assessment of tumor response to therapy. MRI strategies are providing new diagnostic methods of evaluating dynamic pathophysiological processes using time-resolved imaging. These methods include the use of an injected contrast agent that can be visualized as it perfuses and then redistributes within the tissues, much as a radiotracer is sometimes used for nuclear imaging techniques (51,52). The advantages of MRI for this application include the capacity for high temporal and spatial resolution and the ability to combine this component of an examination with other elements of MRI to provide comprehensive evaluation during a single study. This approach is being used to detect and characterize diseases, including neoplasms, inflammation and infection, and fibrotic processes (53–57).

Apoptosis, or programmed cell death, is an elementary process of life that maintains homeostasis. Apoptosis is physiologic and is different from necrosis, which is pathologic, often eliciting an inflammatory reaction. Successful chemotherapy increases apoptosis and enhances the uptake of annexin in the tumor. As the tumor becomes necrotic, the uptake of annexin reduces. PET and SPECT imaging with radiolabeled annexin have been used to assess apoptosis (58–61).

Molecular processes that occur within breast cancer cells can be imaged by using a radiotracer based on amino acid metabolism, such as methionine, which can serve as an indirect measure of protein synthesis, or an amine, such as choline, which can reflect membrane biosynthesis. 11C-methionine and 11C-choline have been used in assessing breast cancer and metastases (62,63). DNA synthesis within breast cancer has been examined with thymidine analogues such as 3'-deoxy-3'-[18F]fluorothymidine (FLT), which has been employed to image intratumoral proliferation (64,65). [15O]H2O PET can also be useful for tumor perfusion (65). These radiotracers can serve to characterize tumor as well as measure therapeutic response.

Estrogen and progestin receptors offer another avenue to image breast cancer. Estrogen receptors (ERs) have been successfully targeted using 16α-[18F]fluoro-17β-estradiol (FES), and early clinical studies have shown them to be as sensitive as FDG imaging and probably more specific (66). In one study of patients with ER-positive breast cancer, the patients also underwent FES PET. PET was able to predict global response to hormonal therapy in human epidermal growth factor receptor 2 (HER2)-negative patients better than pathology (24% vs. 47%) (67,68). Radiolabeled chemotherapeutic agents such as [18F]fluorotamoxifen may be useful to detect ER-positive metastases and can also help in following response to treatment (69). 18F-labeled paclitaxel has also been shown in a mouse model with a breast carcinoma xenograft to predict tumor response to paclitaxel chemotherapy (70).
Peptides that localize to receptors are small-sized molecules and clear rapidly via the kidney, making these peptides attractive imaging probes. Peptides, such as bombesin (BN), labeled with single-photon and dual-photon radionuclides are being studied both for imaging and for therapeutic purposes (71). Gastrin-releasing peptide (GRP) receptors, a subtype of BN receptors, are overexpressed in breast carcinoma. Octreoscan, a single-photon radiotracer somatostatin analogue, has also been used to study breast cancer (72).

Several gene therapy approaches have been considered for the management of breast cancer, such as ablation of oncogenic products, restoration of ER expression, alteration of genes that are involved in apoptosis, and activation of tumor suppressor genes. Reporter genes have been used to study promoter or regulatory elements involved in gene expression conventionally studied by tissue biopsy and immunohistochemistry. Now noninvasive in vivo methods are available in the form of radionuclide techniques or optical imaging. The ability to image deeper structures in larger animal models and in humans is the main advantage of radionuclide technique over optical imaging. For example, 8-[18F] Fluoropenciclovir has been studied as a reporter probe for herpes simplex virus type 1 thymidine kinase (HSV1-tk) gene expression (73). Other possibilities still in early development include specially designed, small radiolabelled antisense oligodeoxynucleotides that may help image gene expression (20). PET imaging with an appropriate PET reporter gene and probe combination may become a powerful tool for molecular imaging of human gene expression (74).

HER2/neu overexpression in breast cancer is associated with poor response to chemotherapy and an unfavorable prognosis. Phase I clinical trials of E1A gene therapy targeting HER2/neu-overexpressed breast cancer has shown increased apoptosis (75). A novel PET radiotracer, (68)Ga-DOTA-F(ab’)(2)-herceptin, has been developed, which has the potential to follow tumor response to the heat shock protein 90 (Hsp90) inhibitor 17-allylamino-17-demethoxygeldanamycin (17AAG) (76). In an animal model, downregulation of HER2 occurs independently of glycolytic response as measured by 18F-FDG-PET. This method could be used to rapidly screen tumor response to a variety of HER2 inhibitors and choose the correct agent within hours rather than wait for weeks or months to determine if there is a response. Multidrug resistance as manifested by MDR1 P-glycoprotein transport activity also has the potential to be studied with SPECT and PET radiotracers (77).

Angiogenesis, the outgrowth of new capillaries, is an important component in tumor formation and spread. Imaging of angiogenesis for diagnosis and following therapy can be performed with PET and SPECT radiotracers targeting integrin, VEGF, and other biomarkers (78). For example, (64)Cu-DOTA-VEGF (121) PET has been successful in small animals to specifically image in vivo expression of VEGF (79). Tumor hypoxia may result in resistance to therapy. [18F]Fluoromisonidazole-3-fluoro-1-(2'-nitro-1'-imidazolyl)-2-propanol ([18F]MISO) and Cu-diacetyl-bis(N4-methylthiosemicarbazone) (Cu-ATSM) radiotracers for PET and blood
oxygen level–dependent (BOLD) MRI are some of the techniques being utilized for in vivo imaging but require further study (80,81).

Finally, nanoparticles technology may be utilized to achieve diagnostic imaging with both PET and MRI and enable combined modality radionuclide therapy with intracellular drug delivery (82). The feasibility of magnetic frequency–activated thermal necrosis with antibody-linked iron oxide nanoparticles conjugated with the single-photon radiotracer Indium-111 has been recently demonstrated in breast cancer xenografts (83).

An understanding of pathology, together with the capabilities of MRI, facilitates the use of a particular MRI feature for the evaluation of a range of diseases that affect different organ systems. The application of MRI time–resolved perfusion imaging for tumor ranges from primary diagnosis to the determination of the degree of tumor neoangiogenesis. The ability to quantify tumor angiogenesis may provide a useful surrogate marker of tumor responsiveness prior to initiating therapy or indicate tumor response even prior to any change in the overall volume of tumor tissue (Fig. 4). MRI may be useful for the specific evaluation of chemotherapeutic antiangiogenic agents such as anti–VEGF (84,85). The possibility of using MRI molecular imaging capabilities for directing therapy is being explored, including the detection of reporter gene transcription, the use of MRI tissue thermometry to guide therapeutic ultrasound for noninvasive tumor ablation, and the use of functional brain MRI to plan the surgical resection of tumors, including metastases to the brain (51,52,86–91).

Most importantly, in a manner analogous to the fusion of PET and CT scanning, the fusion of PET and MRI scanning holds even greater promise in providing a one-stop shop for molecular function and exquisite anatomic detail but with a markedly reduced radiation dose. There have been a number of studies examining the utility of MRI and FDG PET. For example, Walter et al. studied 42 lesions in 40 patients preoperatively (92). The sensitivity of FDG PET was 63%, and the specificity was 91%; the sensitivity of MRI was 89%, and the specificity was 74%. Thus, both modalities can play complementary roles, with overall improved accuracy when used in conjunction. Other papers examining the complementary strengths and weaknesses of both modalities have emerged (93–95). A recent intriguing study in a small group of patients examines the utility of MRI with ultrasmall superparamagnetic iron oxide particles in conjunction with PET, achieving 100% sensitivity and specificity on a per-patient basis for local lymph node staging when both modalities are employed (96).

Further technical development is anticipated to facilitate combining MRI and PET (97). We have already started accumulating evidence of the advantages in this pairing of technologies, as compared with using the combination of CT and PET. Figure 5 is an example from our institution of how PET and MRI data can be combined utilizing special software registration. While the above PET-MRI studies have been performed on separate devices, necessitating moving and scheduling the patients between two physically separate units, a combined PET-MRI scanner in a single gantry is now a reality (45,46,97,98).
OUR VISION

Molecular imaging will continue to have an evolving role in diagnosis, staging, personalization of therapy, monitoring response to therapy, and detecting recurrence of breast carcinoma, both local and widespread. A patient will arrive in the imaging suite and have the entire disease process interrogated on a specially modified PET-MRI system, which will utilize dynamic MRI techniques and MR spectroscopy as well as novel PET radiotracers. Complex processes such as glycolysis, tumor perfusion, cell proliferation, protein synthesis, genetic expression, angiogenesis, and various other profiling will be examined. There is potential for additional information and improvements over current invasive tissue sampling and histopathology, which, for example, may not be truly representative of the heterogeneity present within a primary tumor or metastases.

This process will not simply serve as a baseline to monitor therapy but will help individualize therapy since all the parameters can be entered into a matrix and therapy chosen on the basis of data derived from a cumulative outcomes database, including surgical, chemotherapeutic, radiotherapy, and genetic therapy options. Drug-target matching and optimal dosing can be assessed utilizing the powerful molecular techniques described above. These techniques will also facilitate targeted biopsies, which will improve genomic and proteomic profiling and permit a more tailored radiotherapy. Designer therapeutic and imaging agents may also be synthesized on the basis of a patient’s particular profile.

Figure 5  Breast cancer patient (from Fig. 4) with axial MRI, PET, and CT images through the pelvis. (A) Shows the MRI and PET images before and after deformable registration of the PET to the MRI. (B) The original CT (upper left) and PET-CT (lower left), compared with the MRI (upper right) and deformable registered PET-MRI (lower right). Increased FDG activity in the perirectal space on PET-CT (arrow) is of uncertain origin; however, the activity from the PET scan clearly corresponds to a benign fibroid on the MRI, which is not visible on CT. Benign fibroids may occasionally take up FDG in premenopausal women. The fusion of PET and MRI increased diagnostic accuracy in this patient. Abbreviations: MRI, magnetic resonance imaging; PET, positron emission tomography; CT, computed tomography; FDG, $^{18}$F-fluorodeoxyglucose.
These same methods may then be used to quickly determine treatment response as measured by glycolysis, amino acid transport, tumor perfusion, proliferation, and apoptosis, among other metabolic indices, and also monitor for resistant or mutating clones and for recurrence. Chemotherapeutic and genetic therapies may then be altered or doses optimized on the basis of rapid molecular imaging feedback rather than maximizing therapy based only on toxicity effects. The possibility of therapy using nanoparticle delivery of killing doses of radiotherapy combined with chemotherapy may also be possible. The same nanoparticles could also carry MR and PET imaging tracers.

The concept of the “one magic bullet” for the generic enemy of cancer has fallen by the wayside. A host of ordinary tools optimized for a particular patient and his or her own cancer can be combined in a unique fashion with the help of the “magic window” afforded by molecular imaging. Though defeating the cancer is a lofty goal, these tools may also enable patients to live with cancer and manage the entire disease process holistically in cooperation with their medical caregivers. The true cost-benefit ratios of therapies can then be ascertained. We believe the technology already exists, but further research, social, and financial investment, as well as societal vision is required for this dream to flower into everyday reality.

REFERENCES


INTRODUCTION

Cancer nanotechnology is an interdisciplinary area of research in science, engineering, and medicine with broad applications in molecular profiling and individualized therapy (1–4). The basic rationale is that nanometer-sized particles such as biodegradable micelles, semiconductor quantum dots (QDs), and iron oxide nanocrystals have functional or structural properties that are not available from either molecular or bulk materials (5–15). When linked with biotargeting ligands such as monoclonal antibodies, peptides, or small molecules, these nanoparticles are used to target malignant tumors with high affinity and specificity. In the “mesoscopic” size range of 5 to 100 nm diameter, nanoparticles also have large surface areas and functional groups for conjugating to multiple diagnostic (e.g., optical, radioisotopic,
or magnetic) and therapeutic (e.g., anticancer) agents. The emergence of nanotechnology has opened new opportunities for personalized oncology in which cancer detection, diagnosis, and therapy are tailored to each individual’s molecular profile, and also for predictive oncology in which genetic or molecular information is used to predict tumor development, progression, and clinical outcome.

Recent advances have led to the development of functional nanoparticles that are covalently linked to biological molecules such as peptides, proteins, nucleic acids, or small-molecule ligands (1–4). Medical applications have also appeared, such as the use of superparamagnetic iron oxide nanoparticles as a contrast agent for lymph node prostate cancer detection (16), and the use of polymeric nanoparticles for targeted gene delivery to tumor vasculatures (17). In particular, the U.S. Food and Drug Administration (FDA) recently approved Abraxane\textsuperscript{TM}, an albumin-paclitaxel (Taxol\textsuperscript{TM}) nanoparticle for the treatment of metastatic breast cancer (18). A phase I clinical trial determined that the maximum tolerated dose (MTD) of single-agent albumin-bound paclitaxel every three weeks was 300 mg/m\textsuperscript{2} in patients with solid tumors (breast cancer and melanoma). A second phase I trial demonstrated five patient responses among 39 pretreated patients with advanced solid tumors, including one response in a patient with non–small cell lung cancer (NSCLC), three responses in patients with ovarian cancer, and one in breast cancer. The dose-limiting toxicity was myelosuppression, the MTD was 270 mg/m\textsuperscript{2}, and premedication was not required. Subsequent use of Abraxane in both phase II and phase III trials proved that this new formulation was considerably superior to Taxol. The use of nanoparticles for drug delivery and targeting is one of the most exciting and clinically important applications of cancer nanotechnology.

**CANCER BIOMARKERS AND NANOTECHNOLOGY**

Cancer markers are broadly defined as altered or mutant genes, RNA, proteins, lipids, carbohydrates, small metabolite molecules, and altered expression of those that are correlated with a biological behavior or a clinical outcome (19–22). Most cancer biomarkers are discovered by molecular profiling studies on the basis of an association or correlation between a molecular signature and cancer behavior. In the cases of both breast and prostate cancers, a major progression step is the appearance of so-called “lethal phenotypes” (causing patient death) such as bone metastatic, hormone-independent, and radiation- and chemotherapy-resistant phenotypes. It has been hypothesized that each of these aggressive behaviors or phenotypes could be understood and predicted by a defining set of biomarkers (20). By critically defining the inter-relationships between these biomarkers, it could be possible to diagnose and determine the prognosis of a cancer on the basis of a patient’s molecular profile, leading to personalized and predictive medicine.

One example is the drug trastuzumab (Herceptin\textsuperscript{TM}; Genentech/Roche, California, U.S./Basel, Switzerland), a monoclonal antibody designed to target
amplified and overexpressed ERBB2 (also known as HER2) tyrosine kinase receptor found in only approximately 25% to 30% of breast cancers. FDA approval of trastuzumab was predicated on the availability of a test to detect ERBB2 overexpression. Both an immunohistochemistry (IHC) assay for the expressed protein (HercepTest™; Dako, Glostrup, Denmark) and a nucleic acid–based fluorescence in situ hybridization (FISH) test (PathVysion; Abbott Laboratories, Illinois, U.S.) have been approved as in vitro diagnostics to guide trastuzumab treatment decisions. In another example, the clinical response of lung cancer patients to the epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor gefitinib (Iressa™; AstraZeneca, London, U.K.) is associated with a small number of genetic mutations (23,24). Thus, a molecular diagnostic test could be used to identify patients that are most likely to respond to this drug.

Despite these advances, critical studies that can clearly link biomarkers with cancer behavior remain a significant challenge. One difficulty is that most cancer tumors (especially prostate and breast cancer) are highly heterogeneous, containing a mixture of benign, cancerous, and stromal cells. Current technologies for molecular profiling including reverse-transcriptase polymerase chain reaction (RT-PCR), gene chips, protein chips, two-dimensional gel electrophoresis, biomolecular mass spectrometry (e.g., MALDI-MS, i.e., matrix assisted laser desorption ionization mass spectrometry, ES-MS, i.e., electro spray mass spectroscopy, and SELDI-MS, i.e., surface enhanced laser desorption/ionization) are not designed to handle this type of heterogeneous samples (25,26). Furthermore, a limitation shared by all these technologies is that they require destructive preparation of cells or tissue specimens into a homogeneous solution, leading to a loss of valuable three-dimensional cellular and tissue morphological information associated with the original tumor. The development of nanotechnology, especially bioconjugated nanoparticles, provides an essential link by which biomarkers could be functionally correlated with cancer behavior (Fig. 1).

**NANOTYPING**

Semiconductor QDs are tiny light-emitting particles on the nanometer scale, and are emerging as a new class of fluorescent labels for biology and medicine (8–15). In comparison with organic dyes and fluorescent proteins, QDs have unique optical and electronic properties such as size-tunable light emission, superior signal brightness, resistance to photobleaching, and simultaneous excitation of multiple fluorescence colors (Fig. 2). These properties are most promising for improving the sensitivity and multiplexing capabilities of molecular histopathology and disease diagnosis. Recent advances have led to highly bright and stable QD probes that are well suited for profiling genetic and protein biomarkers in intact cells and clinical tissue specimens (2–4). In contrast to in vivo imaging applications where the potential toxicity of cadmium-containing QDs is a major concern, immunohistological staining is performed on in vitro or ex vivo clinical
patient samples. As a result, the use of multicolor QD probes in IHC is likely one of the most important and clinically relevant applications in the near term (27).

Indeed, significant opportunities exist at the interface between QD nanotechnology and signature biomarkers for cancer diagnosis and individualized therapy. In particular, QD nanoparticle probes have been used to quantify a panel of biomarkers on intact cancer cells and tissue specimens, allowing a correlation of traditional histopathology and molecular signatures for the same material (2–4,27). A single nanoparticle is large enough for conjugation to multiple ligands, leading to enhanced binding affinity and exquisite specificity through a “multivalency” effect (28). These features are especially important for the analysis of cancer biomarkers that are present at low concentrations or in small numbers of cells.

Quantitative biomarker information can be obtained by using a spectrometer attached to the fluorescence microscope. It is however very important to use a common protein such as β actin or glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as an “internal control.” That is, one of the QD–Ab (antibody) conjugates should be designed to measure the product of a housekeeping gene that is expressed at relatively constant levels in all cells. The use of an internal control holds great promise for overcoming a number of major problems in biomarker quantification, such as differences in the probe brightness, variations in probe binding efficiency, uneven light illumination, and detector responses.
The majority of available tumor specimens are archived, formalin-fixed paraffin-embedded (FFPE) tissues that might be several decades old. As the clinical outcomes of these tissues are already known, these specimens are well suited for examining the relationship between molecular profile and clinical outcome in retrospective studies.

It is also critically important to validate the QD staining data with other available techniques. For this purpose, O’Regan and coworkers (2) have obtained QD molecular profiling data from standard human breast cell specimens and have compared the corresponding biomarker data with traditional IHC and FISH techniques. Briefly, slides from formaldehyde-fixed paraffin cell blocks were stained in accordance with standard pathological protocols for three breast cancer biomarkers—estrogen receptor (ER), progesterone receptor (PR), and HER2. This panel of protein biomarkers was selected because of its clinical significance in human breast cancer diagnosis and treatment (29–32). The traditional IHC results were analyzed by two independent observers and scored with a standard scale from 0 (no visible staining in the nucleus or membrane) to

![Cancer Cell](image1)

![QD-Cancer Cell](image2)

Figure 2 Nanotechnology linking biomarkers with cancer behavior and clinical outcome. The schematic diagram shows multiplexed detection and quantification of cancer biomarkers on intact cells or tissues with multicolor nanoparticle probes. (Left) The images show cancer cells labeled with quantum dots, and (right) the drawings suggest how surface and intracellular biomarkers could be quantified by wavelength-resolved spectroscopy or spectral imaging.
3+ (strong and complete membrane or nuclear staining in > 10% of malignant stained cells). For a comparative analysis of QD profiling with traditional IHC, it is necessary to normalize the absolute fluorescence intensities of QD–Ab signals, so that relative percentage values are calculated from the maximum signal strength.

The results reveal that a 3+ score for ER, PR, or HER2 by traditional IHC corresponds to 85% to 100% relative expression of the antigen by QD–Ab measurement, and that 1+ or 2+ scores by traditional IHC correspond to 11% to 48% expression as determined by QD quantification. It should be noted that classification of antigens expressed at low levels (1+ or 2+) is subjective, requiring experience and often resulting in considerable interobserver variations. In contrast, quantitative QD measurements allow accurate determination of tumor antigens at low levels. For example, PR expression in MCF-7 cells and ER expression in BT-474 cells are both classified as 1+ by traditional IHC, but quantitative QD measurements indicate major differences in PR expression (16.8%) and ER expression (47.7%) in these two cell lines. This indicates that the quantitative nature of QD-based molecular profiling could simplify and standardize categorization of antigens that are expressed at low levels. This is of fundamental importance in the management of breast cancer, as the likely benefit of hormonal therapies and trastuzumab depends directly on not just the presence but also the quantity of hormone or HER2 receptors (33–35) (Fig. 3).

Wang and coworkers (27) have developed an integrated image processing and bioinformatics software tool called Q-IHC for quantitative analysis of biomarker expression and distribution in IHC images. In comparison with previous image processing software for automated feature extraction and quantitative

![Figure 3](image-url) Targeting the HER 2/Neu antigen on breast cancer cells by using antibody-conjugated quantum dots. The image shows individual breast cancer cells with nuclear counterstaining.
Cancer Nanotechnology

analysis (36,37), this software system is capable of handling imaging data from both traditional and QD-based IHC. To measure the distribution of labeled antigens, multiple slides of IHC imaging data are acquired to capture selected tissue structures. After image acquisition, an image-processing module carries out automatic boundary identification, semiautomatic image segmentation, and color-based tissue classification on the basis of biomarker staining. Then, an image analysis module quantifies the various biomarker features into numerical values. These values become distinct features and are used for comparison with clinical diagnosis. After validation by a physician, the quantitative data and rules describing biomarker features are stored in a database. This semiautomatic image processing and quantification system is designed to provide molecular profiling data that are more objective, more consistent, and more reproducible than completely manual or automated quantification methods. The software tools process image files from slide scanners in Matlab, which is a collection of various engineering processing tools. In addition, a user-friendly graphical interface allows users to give input and feedback to improve the system quality.

NANOTHERAPEUTICS

Most current anticancer agents do not greatly differentiate between cancerous and normal cells, leading to systemic toxicity and adverse effects. Consequently, systemic applications of these drugs often cause severe side effects in other tissues (such as bone marrow suppression, cardiomyopathy, and neurotoxicity), which greatly limit the maximal allowable dose of the drug. In addition, rapid elimination and widespread distribution into nontargeted organs and tissues require administration of a drug in large quantities, which is not economical and often complicated because of nonspecific toxicity. Nanotechnology offers a more targeted approach and could thus provide significant benefits to cancer patients.

Rapid vascularization in fast-growing cancerous tissues is known to result in leaky, defective architecture and impaired lymphatic drainage. This structure allows an enhanced permeation and retention (EPR) effect (38–42), as a result of which nanoparticles accumulate at the tumor site. For such passive targeting mechanism to work, the size and surface properties of drug delivery nanoparticles must be controlled to avoid uptake by the reticuloendothelial system (RES) (43). To maximize circulation time and targeting ability, the optimal size should be less than 100 nm in diameter and the surface should be hydrophilic to circumvent clearance by macrophages. A hydrophilic surface of the nanoparticles safeguards against plasma protein adsorption, and it can be achieved through hydrophilic polymer coatings such as polyethylene glycol (PEG), poloxamines, poloxamers, polysaccharides or through the use of branched or block amphiphilic copolymers (44–47). The covalent linkage of amphiphilic copolymers (polylactic acid, polycaprolactone, and polycyanonacrylate chemically coupled to PEG) is generally preferred, as it avoids aggregation and ligand desorption when in contact with blood components.
An alternative passive targeting strategy is to utilize the unique tumor environment in a scheme called tumor-activated prodrug therapy. The drug is conjugated to a tumor-specific molecule and remains inactive until it reaches the target (48). Overexpression of the matrix metalloproteinase (MMP), MMP-2, in melanoma has been shown in a number of preclinical as well as clinical investigations. Mansour et al. (49) reported a water-soluble maleimide derivative of doxorubicin (DOX) incorporating an MMP-2 specific peptide sequence (Gly-Pro-Leu-Gly-Ile-Ala-Gly-Gln) that rapidly and selectively binds to the cysteine-34 position of circulating albumin. The albumin-DOX conjugate is efficiently and specifically cleaved by MMP-2, releasing a DOX tetrapeptide (Ile-Ala-Gly-Gln-DOX) and subsequently DOX. pH and redox potential have been also explored as drug-release triggers at the tumor site (50). Another passive targeting method is the direct local delivery of anticancer agents to tumors. This approach has the obvious advantage of excluding the drug from the systemic circulation. However, administration can be highly invasive, as it involves injections or surgical procedures. For some tumors such as lung cancers that are difficult to access, the technique is nearly impossible to use.

Active targeting is usually achieved by conjugating to the nanoparticle a targeting component that provides preferential accumulation of nanoparticles in the tumor-bearing organ, in the tumor itself, individual cancer cells, or intracellular organelles inside cancer cells. This approach is based on specific interactions such as lectin-carbohydrate, ligand-receptor, and antibody-antigen (51). Lectin-carbohydrate is one of the classic examples of targeted drug delivery (52). Lectins are proteins of nonimmunological origin, which are capable of recognizing and binding to glycoproteins expressed on cell surfaces. Lectin interactions with certain carbohydrates are very specific. Carbohydrate moieties can be used to target drug delivery systems to lectins (direct lectin targeting), and lectins can be used as targeting moieties to target cell surface carbohydrates (reverse lectin targeting). However, drug delivery systems based on lectin-carbohydrate have mainly been developed to target whole organs (53), which can pose harm to normal cells. Therefore, in most cases, the targeting moiety is directed toward specific receptors or antigens expressed on the plasma membrane or elsewhere at the tumor site.

The cell surface receptor for folate is inaccessible from the circulation to healthy cells because of its location on the apical membrane of polarized epithelia, but it is overexpressed on the surface of various cancers like ovary, brain, kidney, breast, and lung malignancies (54,55). Surface plasmon resonance studies revealed that folate-conjugated PEGylated cyanoacrylate nanoparticles had a 10-fold higher affinity for the folate receptor than the free folate ones did (56). Folate receptors are often organized in clusters and bind preferably to the multivalent forms of the ligand. Furthermore, confocal microscopy demonstrated selective uptake and endocytosis of folate-conjugated nanoparticles by tumor cells bearing folate receptors. Interest in exploiting folate receptor targeting in cancer therapy and diagnosis has rapidly increased, as attested by many
conjugated systems including proteins, liposomes, imaging agents, and neutron activation compounds (54,55).

For enhanced tumor-specific targeting, the differences between cancerous cells and normal cells may be exploited. By virtue of their small size, nanoparticles entail a high surface area that not only paves a way for more efficient drug release but also a better strategy for functionalization. There is a growing body of knowledge on unique cancer markers due to recent advances in proteomics and genomics. They form the basis of complex interactions between bioconjugated nanoparticles and cancer cells. Carrier design and targeting strategies may vary according to the type, developmental stage, and location of cancer (57). There is much synergy between imaging and nanotechnology in biomedical applications. Many of the principles used to target delivery of drugs to cancer may also be applied to target imaging and diagnostic agents to enhance detection sensitivity in medical imaging. With engineered multifunctional nanoparticles, the full in vivo potential of cancer nanotechnology in targeted drug delivery and imaging can be realized.

**BIONANOINFORMATICS**

For the long-term goal of personalized and predictive oncology, it is important to integrate biotechnology and nanotechnology with information technology (Fig. 4). A major step in this direction is the development of the Cancer Bioinformatics Grid (caBIG), the “World Wide Web for Cancer Research” spearheaded by the U.S. National Cancer Institute (NCI) (58). Through data sharing and standardization, this national initiative aims at improving the infrastructures of health care information technology, to facilitate the process from clinical data

---

**Figure 4** Schematic diagram showing the integration of biotechnology, nanotechnology, and information technology for personalized oncology.
collection to computational data mining, and to accelerate the use of biomolecular markers for personalized treatment. The caBIG organizational structure is divided into working groups, each focused on different aspects of the interoperability problem specific to cancer research. Two overarching work groups, “Architecture” and “Vocabularies and Common Data Elements” have begun to produce detailed specifications that will guide the design of interoperable bioinformatics tools. In practice, caBIG provides a common platform from which to launch even more ambitious integrated solutions. Bioinformatics laboratories can adopt the basic infrastructure, databases, and functionality provided by caBIG working groups to fill gaps in knowledge and to carry out information management responsibilities. In addition to these shared tools and architectures, caBIG is the perfect test bed for a suite of standard tools that can be understood, verified, and used by clinicians around the world. Another major effort is Cancer Bioinformatics Infrastructure Objects (caBIO), an application programming interface (API) that was first developed as Java objects and was later extended to support various programming platforms. These objects represent the most fundamental concepts in bioinformatics research such as genes, ontology, and sequences. Also under development are network programs such as caBIONet based on Microsoft’s .NET technology for data analysis and integration (59).

OUTLOOK
Looking into the future, there are a number of research themes or directions that are particularly promising but require concerted effort for success. The first direction is the design and development of nanoparticles with monofunctions, dual functions, three functions, or multiple functions. For cancer and other medical applications, important functions include imaging (single or dual-modality), therapy (single drug or combination of two or more drugs), and targeting (one or more ligands). With each added function, nanoparticles could be designed to have novel properties and applications. For example, binary nanoparticles with two functions could be developed for molecular imaging, targeted therapy, or for simultaneous imaging and therapy (but without targeting). Bioconjugated QDs with both targeting and imaging functions will be used for targeted tumor imaging and molecular profiling applications. Conversely, ternary nanoparticles with three functions could be designed for simultaneous imaging and therapy with targeting, targeted dual-modality imaging, or for targeted dual-drug therapy. Quaternary nanoparticles with four functions can be conceptualized in future to have the abilities of tumor targeting, dual-drug therapy, and imaging. The second direction is nanoparticle molecular profiling (nanotyping) for clinical oncology; i.e., the use of bioconjugated nanoparticle probes to predict cancer behavior, clinical outcome, treatment response, and individualize therapy. This should start with retrospective studies of archived specimens because the patient outcome is already known for these specimens.
Cancer Nanotechnology

The key hypotheses to be tested are that nanotyping a panel of tumor markers will allow more accurate correlations than single tumor markers, and that the combination of nanotyping tumor gene expression and host stroma are both important in defining the aggressive phenotypes of cancer as well as determining the response of early-stage disease to treatment (chemotherapy, radiation, or surgery). The third important direction is to study nanoparticle distribution, excretion, metabolism, and pharmacodynamics in in vivo animal models. These investigations will be very important in the development of nanoparticles for clinical applications in cancer imaging or therapy.

In conclusion, nanotechnology is emerging as an enabling technology for personalized oncology in which cancer detection, diagnosis, and therapy are tailored to each individual’s tumor molecular profile and also for predictive oncology in which genetic or molecular markers are used to predict disease development, progression, and clinical outcomes (60,61).

ACKNOWLEDGEMENTS

We are grateful to Dr. Leland Chung, Dr. Ruth O’Regan, and Dr. Dong Shin for insightful discussions. Our research in cancer nanotechnology was supported by National Institutes of Health (NIH) grants (R01 CA108468-01 and U54CA119338) and the Georgia Cancer Coalition Distinguished Cancer Scholars Program (to May Dongmei Wang and Shuming Nie).

REFERENCES

32. Elledge RM, Green S, Pugh R, et al. Estrogen receptor (ER) and progesterone receptor (PR), by ligand-binding assay compared with ER, PR and pS2, by


Index

ABCB1 polymorphisms, 19–20
Abraxane™, 310
Acute myeloid leukemic (AML) cells, 271
Adenosine triphosphate (ATP), 52
Adjuvant chemotherapy, 100–101, 207
Adjuvant! Online, 187, 189, 205, 207
Adjuvant on Line (AOL), vs. RS assay, 103–104
AdnaGen system, 157
Adriamycin. See Doxorubicin
Advancing Clinico Genomic Trials (ACGT), 207
16α-[^18F] fluoro-17β-estradiol (FES), 298
Affymetrix GeneChip, 250, 266
Affymetrix microarray, 239
AIB1, 171
Akt, 173, 277
Aldefluor®, 270
Aldehyde dehydrogenase (ALDH), 270–271
ALDH. See Aldehyde dehydrogenase (ALDH)
17-allylamo-17-demethoxygeldanamycin (17AAG), 299
American College of Surgeons Oncology Group (ACOSOG), 234
American Joint Committee on Cancer (AJCC) classification system, 159–160
Aminoglutethimide, 174
Amonafide, 17
Amphiphilic copolymers, covalent linkage of, 315
Amplicons, 244
Anastrozole, 32, 174
chemical structure of, 175
vs. tamoxifen, 177
Angiogenesis
PET and SPECT imaging of, 299
Animal model, 299
Anthracyclines
CBR and, 28
oxidative stress and, 28–29
Antiaapoptosis genes, 201, 202
Anticytokeratin antibodies, 153, 154
Antiestrogens, 168–174
Anti-VEGF, 300
API. See Application programming interface (API)
Apoptosis, 298
Application programming interface (API), 318
Aromatase inhibitors (AI), 32, 137
analysis of gene expression with, 177–178
chemical structures of, 175
mechanisms of resistance, 175–176
steroidal and nonsteroidal, 174
vs. tamoxifen, 176–177
Array CGH, 250
ATAC trial, 175–177
AutoMACSTM system, 153–155
Automatic magnetic cell sorting system.
See AutoMACSTM system
Avidin-biotin complex peroxidase
technique, 154

BAC. See Bacterial artificial
cromosomes (BAC)
Bacterial artificial chromosomes (BAC), 64–65
β-catenin, 273
Bcl-2 tumor supressor, 248
BCRP1 (breast cancer resistance
protein 1), 268
17β-estradiol (E2), 168, 201
Bioinformatics laboratories, 318
Biomarkers, 51. See also Cancer
biomarkers
Bionanoinformatics, 317–318
Biopsies, gene expression profiling in, 91–92
Biotinylated random nonamers, 246
Blood collection tubes, 216–217
Blood tissue, 218
Bmi-1, 271–272
BMI-1 oncogenic pathway, 202
BOLD (blood oxygen level-dependent)
MRI, 300
Bone marrow
cancer stem cells in, 152
cytokeratin positive tumor cells in, 206
micrometastasis, 154
tumor cells, 154
Bone mineral density, 178
Bone scanning, 291, 292
BRCA1, 5, 50–51, 266, 267
BRCA2, 5
Breast cancer
genic susceptibility, 5
luminal, 198
molecular classification of, 3, 198–199
patients information
storage and protection of, 231–232
treatment with genetic
polymorphism, 267
Breast cancer cells
circulating. See Circulating tumor cells
subgroups of, 202–204
Breast Cancer Intergroup, 186, 188, 190
Breast International Group (BIG) 1-98
study, 175
Breast MRI, 295–296
Breast tumor model, 269
caBIG. See Cancer Bioinformatics Grid
caBIG)
ciaBIONet, 318
Cancer Bioinformatics Infrastructure
Objects (caBIO) API, 318
Cancer biomarkers, 310–311
Cancer Biomedical Informatics Grid
(cabIG), 207
development of, 317
organizational structure of, 318
Cancer cells vs. normal cells, 317
Cancer metastasis, 278
Cancer Methylation Panel, 250, 251
Cancerous tissues, vascularization in, 315
Cancer stem cells. See Stem cells
Cancer Trials Support Unit, 192
CapSure HS LCM Cap, 223, 224
Carbohydrate moieties, 316
Carbonyl reductase (CBR), 28
Carcinogenesis, 269
CAT. See Catalase
Catalase (CAT), 29
11C-choline, 298
CD24− breast cancer cells, 202
CD44+ breast cancer cells, 202
CD44+CD24−/low breast cancer cells,
265, 266, 270
CD44 expression, 270
cDNA synthesis, 246, 247
CellProfileTM kit, 153
CellSearchTM system, 153
for prognostic and predictive value of
CTCs, 158
CellSpotterTM Analyzer, 153
c-erbB-2, 161
CGH. See Comparative genomic
hybridization (CGH)
Chemotherapeutic agents, 263
Index

Chemotherapy, 159
  adjuvant. See Adjuvant chemotherapy
cytotoxic, 161
evaluation of marker, 186
neoadjuvant. See Neoadjuvant chemotherapy
residual clinical risk and, 102–103
Chromatin immunoprecipitation (ChIP), 53
Chromogenic in situ hybridization (CISH), 139
Circulating female hormone, 168
Circulating tumor cells (CTCs)
  baseline and follow-up, 159
CD326-positive, 154, 155
clinical application of, 160–161
detection, 153–157
  clinical benefits of, 161
  future uses of, 162
  immunostains of, 154
  magnetic labeling and separation of, 219–220
  prognostic value of, 158
  for staging of MBC, 159–160
CISH. See Chromogenic in situ hybridization (CISH)
Classification of breast tumors, 250
Clinical decision making, 190
Clinical Laboratory Improvement Amendments (CLIA), 190
Clinical trials
  for antiestrogens, 168
designing, 190–192
  protocols, 214, 228
  for treatment with MAPK-specific inhibitors, 172
Clinicogenomic models, 207
Cloned probe arrays, in CGH, 65–66
Cluster analysis, 248
11C-methionine, 298
CNA, Copy number abnormalities (CNA)
CNB. See Core needle biopsy (CNB)
Coexpression modules, 202
Combined imaging, with FDG-PET and MRI, 297
Common Rule, 232
Comparative genomic hybridization (CGH), 62, 249
BAC, 64–65

[Comparative genomic hybridization (CGH)]
  cloned probe arrays in, 65–66
  DNA in, 69–71
  metaphase chromosomes in, 65
  MIP, 66–69
  oligonucleotide-based arrays, 65–66
  Complete pathological response (pCR), 101–102
Computed tomography (CT) scanning, 294
Contingency plans, 214
Copy number abnormalities (CNA), 61, 249
  assessment by CGH. See Comparative genomic hybridization (CGH)
  gene deregulation by, 62–63
Coregulators, 171
Core needle biopsy (CNB), 83, 238–239
Cox model, 245
Cox regression, 159
CpG dinucleotides, 45–46
Cryopreservation techniques, 235–237
CTCs. See Circulating tumor cells (CTCs)
Cu-ATSM, 299
Cu-DOTA-VEGF, 299
CXCR4 chemokine receptor, 278, 280
Cycling cells, 267, 282
Cyclophosphamide
  CYP3A4 in, 27
  CYP2B6 in, 26–27
  and GST, 27–28
CYP. See Cytochrome p450s (CYP)
CYP3A4, in cyclophosphamide, 27
CYP1B1, 267
CYP2B6, in cyclophosphamide, 26–27
CYP2C8
  in paclitaxel, 31
  CYP2D6
  genotyping, 170
  in TAM, 32
Cytochrome p450s (CYP), 267
Cytology, 51–52
Cytotoxic agents, in neoadjuvant chemotherapy, 114–115
Cytotoxic therapies, 109

Dandelion hypothesis, 260
DASL assay, 246
Database information, and patient identifiers, 232
Daunorubicin, 28
DCIS. See Ductal carcinoma in situ (DCIS)
Deacetylation, of histones, 47
Decision aids, 187
Decision boards, 187
DES. See Diethylstilbestrol (DES)
DFS. See Disease-free survival
3D GRE technique, 296–297
Diethylstilbestrol (DES), vs. tamoxifen, 168
Dihydropyrimidine dehydrogenase (DPD), 18–19
Discovery-based research, 185
Disease-free survival (DFS), 32–36, 109
defined, 192
GST in, 33
MnSOD, 36
MTHFR in, 33
SULT1A1, 36
UGT2B15, 36
Disseminated tumor cells (DTC), 206
Distant recurrence-free survival (DRFS) defined, 192
DNA, in CGH, 69–71
DNA hypermethylation, 279
DNA isolation, 225–226
DNA methylation, 45
DNMT in, 46–47
inhibitors, 53–54, 55
DNA methyltransferases (DNMT), 46–47
DNA microarray analysis, 107–108
DNA synthesis, 298
4',6-diamidino-2-phenylindole (DAPI), 153
Docetaxel, 19–20
Doxorubicin (DOX), 28, 316
DPD. See Dihydropyrimidine dehydrogenase (DPD)
Drosophila, 276
Drugs
chemotherapy, 159
development
chemotherapy, 159
for CTC, 161
predictive markers in, 141–143
DSL ligands, 276
Ductal carcinoma in situ (DCIS), 174
E2. See 17β-estradiol (E2)
Eastern Cooperative Oncology Group (ECOG), 192
EasySep™ system, 155–157
EGFR. See Epidermal growth factor receptor (EGFR)
Embryogenesis, 269
Embryonic stem cells, 271, 279
Endocrine therapy, 167–168
with antiestrogens, 168–174
with aromatase inhibitors, 174–178
Endothelial cells, 279, 280
magnetic labeling and separation of, 219–220
Endoxifen. See 4-hydroxy-N-desmethyl tamoxifen
Enzymes, metabolic, 267
EpCAM. See Epithelial cell adhesion molecule-1 (EpCAM)
Epidermal growth factor receptor (EGFR), 140–141, 171, 172
Epigenetics
DNA methylation. See DNA methylation
histones. See Histones
Epithelial cell adhesion molecule-1 (EpCAM), 153
Epithelial stem cells, 263
ER. See Estrogen receptors (ER)
ERα, 136–138, 168
gene transcription, 169
ERα-positive breast cancer, 168, 173
ERα-positive cells, 269
ERBB2, 62, 311
ER-negative breast cancer, 168, 170
ER-negative tumor cells, 269–270
ER-positive breast cancer, 187, 246
ER-positive tumors, 198, 245
Estrogen receptors (ER), 51, 108, 298
ERα, 136–138, 168
gene transcription, 169
gene expression of, 55–56
mutations, 170
European Human Frozen Tumor Tissue Bank, 234
European Union, 207
Exemestane, 174
centralized structure of, 175
ExtractSure device, 223, 224

FDG PET, 293–294
FES. See 16α-[18F] fluoro-17β-estradiol (FES)
FES PET, 298
[18F]fluorodeoxyglucose (FDG), 292
[18F]fluorotamoxifen, 298
FFPE. See Formalin-fixed paraffin-embedded (FFPE)
FFPE tissues, 313
Fibroblasts, 280
Fibromas, 297
Ficoll-Paque™, density gradient procedure, 156
Fine needle aspiration (FNA), 83, 238
Fixed tissue, 218
[18F]-labeled paclitaxel, 298
Flow cytometry, 264
Fluorescence-activated cell sorter (FACS), 270, 271
Fluorescence microscope, 312
Fluorescent in situ hybridization (FISH), 311
for HER2/neu testing, 138–139
Fluorodeoxyglucose, 293
5-Fluorouracil (5FU)
and DPD, 18–19
DPD and, 30
MTHFR and, 30, 31
TS and, 31
TS inhibition by, 19
[18F]MISO ([18F]fluoromisonidazole-3-fluoro-1-(2’-nitro-1’-imidazolyl)-2-propanol), 299
FNA. See Fine needle aspiration (FNA)
Folate receptors, 316
Food and Drug Administration (FDA), 153, 174, 246, 311, 3111
Formalin, 215, 218
cross-linking effects of, 237
Formalin-fixed paraffin-embedded (FFPE) tissue extraction of RNA from tissue samples of, 98–99, 244–245
RNA-based expression profiling of, 246
Forward genetics approach, 16
FOXO regulatory proteins, 277
Freeze-thawing cycles, 235
Fulvestrant, 174
Ga-DOTA-F(ab')2-herceptin, 299
Gamma rays, 292
Gamma secretase inhibitor (GSI), 276
Gefitinib, 172, 311
Gene expression analysis, 117–178, 244–245
See also Gene expression profiling 70-gene expression, 108, 124–126, 205.
for MINDACT trial, 130–132
prognostic value of, 126–127
in RASTER trial, 130
in TRANSBIG study, 127–129
Gene expression analysis, 177–178, 244–245
Gene expression Grade Index (GGI), 200–201
Gene expression profiling, 79–81, 97–98, 198–199, 246, 270
and chemotherapy, 100–101, 111–114
in neoadjuvant setting, 101–102
for classification of breast cancer, 198–199
and docetaxel, 112–113
oncogenic pathway analysis by, 116–118
pCR prediction, 101–102
RNA expression, 98–100
RS in. See Recurrence score (RS) assay RT-PCR for, 99
Gene expression signatures, 265–267. See also Gene expression profiling for improving prognosis, 199–202
integration with, 202–204
Gene methylation
BRCA1, 50–51
estrogen receptors (ER), 51
RASSF1A, 49–50
21-gene profile. See Oncotype DX™
76-gene profile, 129
Genetic polymorphisms, 266–267
Genetic susceptibility, 5
Gene transcription, 19
Genomic alterations
    prognostic value of, 158
Genomic Health, 190, 191
Genomic Health, Inc. (GHI), 246
Genomic hypomethylation, 279
Genomic predictors, 204–205
Germline polymorphisms, 267
Gestation, 269
GHI. See Genomic Health, Inc. (GHI)
Gliomas, 268
Glutathione-S-transferases (GST), 16, 267
    and cyclophosphamide, 27–28
    in DFS, 33
Glycophorin A, 156
Graphical interface, user-friendly, 315
GRP (gastrin-releasing peptide)
    receptors, 299
GSK-3β, 273
Harmonization, 214
HAT. See Histone acetyltransferases
    (HAT)
Health Insurance Portability and
    Accountability Act, 232
Hedgehog (HH), 265
    ligands, 276–277
    signaling pathway, 276–277
    transcription factors, Gli1 and
    Gli2, 277
H & E (hematoxylin & eosin), 227
Hematopoietic cells, human, 156
Heparin-coated syringes, 215
HER2. See Human epidermal growth
    factor receptor 2 (HER2)
HER2+, 198
HER3, 171
HER4, 171
HER2/neu, 138–139, 171–172
Heterochromatin, 47
HH. See Hedgehog (HH)
Histone acetyltransferases (HAT), 47
Histone deacetylases (HDAC), 55
Histones
    deacetylation, 47
    HAT, 47
    HDAC. See Histone deacetylases
        (HDAC)
    methylation of, 278
    modifiers, 48
    posttranslational modification of, 47
    ChIP for, 53
H3K9 repressors, 278
Homogenization, 224
Hormonal therapy, 191
    adjuvant, 187
Hormone receptors, 161
Hormone replacement therapy
    (HRT), 173
Human epidermal growth factor receptor
    2 (HER2). See also HER2/neu
    gene amplification in CTCs, 158
Hybridization. See Comparative genomic
    hybridization (CGH)
4-hydroxy-N-desmethyl tamoxifen,
    20, 170
Hypothesis-based research, 185
IHC. See Immunohistochemistry (IHC)
Illumina, Inc., 249
Illumina BeadArray, 250
Illumina microarrays, 249
Image file processing, 315
Immunohistochemistry (IHC), 135–136,
    227, 311, 313
for ERz, 136–138
for HER2/neu, 138
for PR, 136–138
Immunohistological staining, 311–312
Indium-111, 300
In situ hybridization (ISH), 227
International Stage IV Stratification
    Study, 160
International Union Against Cancer
    (UICC), 159
Invasiveness gene signature (IGS), 202
Iressa. See Gefitinib
Isoniazid, 16–17
Java objects, 318

Kaiser Permanente health care system, 188
Kaplan-Meier estimator, 245
Krishnamurthy, Savitri, 154

Laser capture microdissection (LCM), 2–3, 221–224, 239
Lauterbur, Paul, 294
LCM. See Laser capture microdissection (LCM)
Lectin-carbohydrate, 316
Lectins, 316
Lethal phenotypes, 310
Letrozole, 32, 174
  chemical structure of, 175
  vs. tamoxifen, 176, 177
Ligand-binding domain (LBD), 170
Liver tumors, 297
Lobular carcinoma in situ (LCIS), 174
Loss of heterozygosity (LOH), 66, 250
Luminal epithelial cells, genes of, 198
Lymph nodes, 297
  microarrays for, 3–5
  recurrence of breast cancer in, 245

Macromolecular extraction, 237, 239
Magnetic labeling and separation, of tumor tissue, 219–220
MammaPrint. See 70-gene expression
Mammary gland, 264
Mammography, 292
Manganese superoxide dismutase (MnSOD), 29
  in DFS, 36
MAPK. See Mitogen-activated protein kinase (MAPK)
MassARRAY, 52
Matlab slide scanners, 315
Matrix metalloproteinase (MMP), 316
MBC. See Metastatic breast cancer (MBC)
MCF-7 breast cancer cells, 170, 172
Medicare, 190
Mercaptopurine, 17
Metaphase chromosomes, in CGH, 65
Metastatic breast cancer (MBC)
  development of, 151
  diagnosis of, 158–159
  stage IV, 160
  HER2-positive, 177, 178
  staging, 159–160
Methotrexate (MTX), 18, 30, 31
Methylation
  detection of, 51–53
  of DNA. See DNA methylation
  of genes. See Gene methylation
  profiling, 250–251
  therapy, 53–55
Methylenetetrahydrofolate reductase (MTHFR), 18
  in DFS, 33
  and 5-Fluorouracil (5FU), 30, 31
Microarray for Node-Negative Disease may Avoid Chemotherapy (MINDACT) trial, 130–132, 205
Microarray Quality Control Project (MAQC), 82
Microarrays
  for lymph node status, 3–5
  with OCT, 236
  with paraffin-embedded tissue, 236–237
  RNA expression, 2, 4
  with RNA stabilization buffers, 237–238
  technical reproducibility of, 82–83
  using CNB, 238–239
  using FNA, 238
  using LCM, 239
Micrometastatic carcinoma, 154
Micrometastatic cells, 206
microRNAs, 206
Microsoft Excel, 228
MIP. See Molecular inversion probe (MIP)
mirNA profiling, 248–249
Mitogen-activated protein kinase (MAPK), 172
MMP. See Matrix metalloproteinase (MMP)
MMP-2, 316
MnSOD. See Manganese superoxide dismutase (MnSOD)

Molecular classification, of breast cancer, 3

Molecular diagnostic test, 186, 188, 190, 311

Molecular epidemiologic studies, 32–36

Molecular imaging, 291–292. See also Positron emission tomography (PET)

Molecular inversion probe (MIP), 66–69

Molecular markers, 79, 160, 188–190

defined, 185–186
development process of, 81–82
inin stem cells, 265, 273

Molecular profiling
defined, 185
developmental technologies for, 311

Molecular signatures, 186

Monoclonal antibodies, 153, 156

Monocytes, 219

MPO. See Myeloperoxidase (MPO)

MRI, 294–297

applications for, 297

mRNAs, 245

MTA-1 (Manual Tissue Arrayer), 226

MTHFR. See Methyltetrahydrofolate reductase (MTHFR)

mTOR (mammalian targets of rapamycin), 276

MTX. See Methotrexate

Mutations

estrogen receptors, 170

Myeloperoxidase (MPO), 29

N-acetyltransferases, 16, 267

Nanoparticles. See also Nanotechnology

bioconjugated, 311, 317
design and development of, 318
molecular profiling. See Nanotyping

Nanotechnology, 311

for cancer detection, diagnosis, and therapy, 319

QD, 312

Nanotherapeutics, 315–317

Nanotyping, 311–315, 318–319

National Cancer Institute (NCI), 186, 207, 317

National Institutes of Health (NIH), 199, 200
cancer guidelines, 205

National Surgical Adjuvant Breast and Bowel Project (NSABP), 99–100, 188, 190, 192, 245

N-CoR, 171

Neoadjuvant chemotherapy, 5–7, 78, 109–111
cytotoxic agents in, 114–115
predictive markers of response to, 101–102
response predictors of, 78–79
use of, 293

Netherlands Cancer Institute (NKI), 199

.NET technology, 318

Noncanonical pathways. See Wnt signaling

Nonnucleoside inhibitors, 53, 54

Non-small cell lung cancer (NSCLC), 310

Nontumorigenic cells, 265

North American Breast Intergroup, 192

Notch receptors, 276

Notch signaling pathway, 264, 276

Nottingham Prognostic Index, 205

NSABP. See National Surgical Adjuvant Breast and Bowel Project (NSABP)

Nucleic acids

amplification, 238
collection, 215

eXtraction, 237

profiling, 251

Nucleoside inhibitors, 53, 54

NuGEN Technologies, Inc., 249

OCT. See Optimal cutting temperature (OCT)

OCT-4, 271

Octreoscan, 299

[15O]H2O PET, 292

4OHT (4-hydroxytamoxifen), 171

Oligonucleotide-based arrays, in CGH, 65–66
Index

Oligonucleotides, 247
Oncogenes, 62, 251
Oncogenic pathways, deregulation of, 116–118
cytotoxic agents, 118
Oncologists, 207
Onctype DX assay, 188, 190, 192, 194, 200, 205, 246
Optimal cutting temperature (OCT), 218, 236

p53
signature genes, 201
tumor suppressor, 248
PACCT. See Program for the assessment of clinical cancer tests (PACCT)
Paclitaxel, 19–20, 267
CYP2C8 in, 31
Paget, Stephen, 278
Papanicolaou (Pap)-stained cytospin, 154
Paraffin, 222, 227
Paraffin-embedded tissue, 237–238
Paraffin wax, discovery of, 237
Partial signatures, 204
Pathologic report, 234
Pathologist, 232, 233
Patient identifiers, 232
PAXgene, 220–221
p21Cip1/Waf1, 277
PCR. See Polymerase chain reaction (PCR); Polymerase chain reaction in MIP
pCR. See Complete pathological response (pCR)
Peptides, 299
Peripheral blood
cancer stem cells in, 152
extraction of RNA from, 220–221
isolation of monocytes in, 219
separation from mononuclear cells, 156
Peripheral blood mononuclear cells (PBMC), 219
PET. See Positron emission tomography (PET)
PET-MRI scanner, 300
PET scanners, 297
Pharmacodynamics, 13
Pharmacoepidemiology, 14
Pharmacogenetics, 14, 206
molecular epidemiologic studies, 32–36
Pharmacogenomic predictors, 79–80
published studies, 86–90
validation studies of, 84–85
Pharmacokinetics, 13
Pharmacology, 13–14
Phosphatase and tensin homolog deleted on chromosome ten (PTEN), 265, 277
Phosphate buffered saline (PBS), 215
Phosphatidylinositol-3-OH kinase (PI3K), 173, 277
Polycomb group (PcG) proteins, 46, 261, 279
Polymerase chain reaction (PCR), 236
detection of tumor cells, 158
in MIP, 67
Positron emission tomography (PET), 206, 292
Posttranslational modification, of histones, 47
PR. See Progesterone receptor (PR)
Predictive markers
defined, 136
in drug development, 141–143
ERz, 136–138
HER2/neu, 138–139
PR, 136–138
Preoperative chemotherapy. See Neoadjuvant chemotherapy
Preoperative systemic chemotherapy (PST). See Neoadjuvant chemotherapy
Preregistration, 191
Preservatives, 214, 215
Primary study group, 192
Privacy rights, 231
Progenitor cells, mutated, 262, 266
Progesterone receptor (PR), 136–138, 176–177
Prognosis, in breast cancer improving, 199–202
Prognostic marker, 199
Program for the assessment of clinical cancer tests (PACCT), 186
Proliferation genes, 201, 202, 204
### Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate cancer, RNA signatures in</td>
<td>248</td>
</tr>
<tr>
<td>Proteinase digestion buffers</td>
<td>237</td>
</tr>
<tr>
<td>Protein biomarkers</td>
<td>313. <em>See also</em> Cancer biomarkers</td>
</tr>
<tr>
<td>Protein collection</td>
<td>215</td>
</tr>
<tr>
<td>Protein isolation</td>
<td>225–226</td>
</tr>
<tr>
<td>Proteomics</td>
<td>206</td>
</tr>
<tr>
<td>PTEN. <em>See</em> Phosphatase and tensin homolog deleted on chromosome ten (PTEN)</td>
<td></td>
</tr>
<tr>
<td>PTEN pathway</td>
<td>277–278</td>
</tr>
<tr>
<td>Pyrosequencing</td>
<td>52–53</td>
</tr>
<tr>
<td>QD-Ab</td>
<td>312</td>
</tr>
<tr>
<td>QDs. <em>See</em> Quantum dots (QDs)</td>
<td></td>
</tr>
<tr>
<td>Q-IHC tools</td>
<td>314</td>
</tr>
<tr>
<td>qRT-PCR (realtime quantitative-PCR) for screening tumor samples for miRNA expression</td>
<td>248</td>
</tr>
<tr>
<td>Quantification system</td>
<td>315</td>
</tr>
<tr>
<td>Quantum dots (QDs)</td>
<td>312</td>
</tr>
<tr>
<td>nanotechnology</td>
<td>312</td>
</tr>
<tr>
<td>semiconductor</td>
<td>311</td>
</tr>
<tr>
<td>Radiation, resistance to (Radioresistance)</td>
<td>266</td>
</tr>
<tr>
<td>Radionuclide techniques</td>
<td>292–294</td>
</tr>
<tr>
<td>vs. optical imaging</td>
<td>299</td>
</tr>
<tr>
<td>RAKE (RNA-primed array-based Klenow extension) analysis study</td>
<td>249</td>
</tr>
<tr>
<td>Randomization</td>
<td>192</td>
</tr>
<tr>
<td>Rapamycin, molecular targets of. <em>See</em> MTOR (mammalian targets of rapamycin)</td>
<td></td>
</tr>
<tr>
<td>RASSF1A</td>
<td>49–50</td>
</tr>
<tr>
<td>RASTER trial</td>
<td>130</td>
</tr>
<tr>
<td>Rationally designed inhibitors</td>
<td>53, 54, 55</td>
</tr>
<tr>
<td>Reactive oxygen species (ROS)</td>
<td>27</td>
</tr>
<tr>
<td>and mitochondrial permeability</td>
<td>29</td>
</tr>
<tr>
<td>Recurrence score (RS)</td>
<td>100, 200, 245</td>
</tr>
<tr>
<td>and adjuvant chemotherapy, 100–101 clinical utility of</td>
<td>103–104</td>
</tr>
<tr>
<td>GHI, 246</td>
<td></td>
</tr>
<tr>
<td>for node positive patients</td>
<td>104</td>
</tr>
<tr>
<td>[Recurrence score (RS)] and pCR</td>
<td>101</td>
</tr>
<tr>
<td>to predict recurrence of breast cancer</td>
<td>245</td>
</tr>
<tr>
<td>vs. AOL</td>
<td>103–104</td>
</tr>
<tr>
<td>REMARK guidelines</td>
<td>186</td>
</tr>
<tr>
<td>Reporter genes</td>
<td>299</td>
</tr>
<tr>
<td>Response predictors, of neoadjuvant chemotherapy</td>
<td>5–7, 78–79</td>
</tr>
<tr>
<td>molecular markers as</td>
<td>79</td>
</tr>
<tr>
<td>Reticuloendothelial system (RES)</td>
<td>315</td>
</tr>
<tr>
<td>Reverse transcriptase-polymerase chain reaction (RT-PCR)</td>
<td>99, 244</td>
</tr>
<tr>
<td>assay-based evaluation of ER expression</td>
<td>189</td>
</tr>
<tr>
<td>and gene signature</td>
<td>200</td>
</tr>
<tr>
<td>RNA amplification</td>
<td>237–238</td>
</tr>
<tr>
<td>degradation</td>
<td>235, 247</td>
</tr>
<tr>
<td>expression microarray</td>
<td>2, 4</td>
</tr>
<tr>
<td>hybridization</td>
<td>238, 239</td>
</tr>
<tr>
<td>isolation</td>
<td>224–225</td>
</tr>
<tr>
<td>RNAlater</td>
<td>218, 236</td>
</tr>
<tr>
<td>RNA stabilization buffers</td>
<td>236–237</td>
</tr>
<tr>
<td>Rosette-Sep™ method</td>
<td>156</td>
</tr>
<tr>
<td>Rotterdam group</td>
<td>199</td>
</tr>
<tr>
<td>RPMI (Roswell Park Memorial Institute) media</td>
<td>218</td>
</tr>
<tr>
<td>RS. <em>See</em> Recurrence score (RS) assay</td>
<td></td>
</tr>
<tr>
<td>RT-PCR assay</td>
<td>245</td>
</tr>
<tr>
<td>Sample handling</td>
<td>237</td>
</tr>
<tr>
<td>Selective ER downregulators (SERDs)</td>
<td>173, 174</td>
</tr>
<tr>
<td>Selective ER modulators (SERMs)</td>
<td>173</td>
</tr>
<tr>
<td>Semiautomatic image processing</td>
<td>315</td>
</tr>
<tr>
<td>Single nucleotide polymorphism (SNP)</td>
<td>14–15, 249</td>
</tr>
<tr>
<td>profiling</td>
<td>250</td>
</tr>
<tr>
<td>Single-photon imaging</td>
<td>292</td>
</tr>
<tr>
<td>SMRT</td>
<td>171</td>
</tr>
<tr>
<td>SNP. <em>See</em> Single nucleotide polymorphism (SNP)</td>
<td></td>
</tr>
<tr>
<td>Sonic hedgehog (SHH)</td>
<td>276, 277</td>
</tr>
<tr>
<td>SRC-1</td>
<td>171</td>
</tr>
<tr>
<td>St. Gallen classification</td>
<td>246</td>
</tr>
<tr>
<td>St. Gallen criteria</td>
<td>199, 200</td>
</tr>
</tbody>
</table>
St. Gallen guidelines, 205
Standard operating procedures (SOPs)
for collection of tissue samples, 234
for RNA degradation, 235
STAR (Study of Tamoxifen and
Raloxifene) trial, 173
STAT3 transcription factor, 276
Stem cells
 Stem cell lines, derivation of, 266
 Stem cell niche, 261, 263, 268
 molecular targets of, 278–280
 Stem cell regulation, 280
 Stem cells
 cellular proliferation, 261
 clinical research on, 264–265
differentiation of, 261, 262
division of, 276
 drug resistance of, 268
 embryonic, 271, 279
 functional targets, 269–273
gene expression profiling of, 270
 hematopoietic, 278, 280
 hypothesis, 260–264, 268
 intestinal, 277
 leukemic, 278
 mesenchymal, 280
 molecular markers, 265, 273
 mutations, 263, 270, 278
 nontumorigenic daughter cells of, 261
 Stem cell sanctuaries, 267–269
 Stem cell survival, neural, 279, 280
 Stochastic model, 260, 261
SULT1A1
 in DFS, 36
 in TAM, 32
 Superoxide dismutases (SODs), 267
 Swiss Institute of Bioinformatics, 202

TAM. See Tamoxifen (TAM)
Tamoxifen (TAM), 20, 31–32, 56, 137
 adjuvant, 169
 CYP2D6 in, 32
 metabolizers of, 170
 SULT1A1 in, 32
 UGT2B15 in, 32
 vs. aromatase inhibitors, 176–177
 vs. DES, 168
 TaqMan™ array, 248–249
Taxanes, 19–20
 CYP2C8 in, 31
 Thiopurine methyltransferase (TPMT), 17–18
 Thymidylate synthase (TS), 19
 Tissue
 acquisition of, 232–235
 biopsies, 218
 collection, 214–218
 control blocks, 228
defined, 213
 processing, 218
 procurement, 214
 Tissue banking, 231–232
 Tissue microarray (TMA), 226–227
 blocks, 228, 229, 230
 mapping, 227, 228
 TMA. See Tissue microarray (TMA)
 TRANSBIG study, 127–129, 199–200
 Transcriptional regulation, 277
 Trastuzumab, 138, 172, 310–311
 Trial Assigning Individualized Options for
 Treatment (TAILORx) trial, 178, 186
designing of, 191
 recurrence score ranges in, 192–194
 TRIZOL method, 224–230
 TS. See Thymidylate synthase (TS)
 Tumor-activated prodrug therapy, 316
 Tumor growth factor-β (TGF-β), 280
 Tumor hypoxia, 299
 Tumorigenesis, 250, 273, 274
 Tumors
 detection of, 158
 invasion, 202
 regression, 260
 size and growth, 238, 239
 Tumor suppressor genes, 251
 Tumor tissue. See Tissue
 UGT2B15
 in DFS, 36
 in TAM, 32
 Ultrasound, 292
### Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacutainers, use of</td>
<td>215</td>
</tr>
<tr>
<td>Vascular endothelial growth factor (VEGF)</td>
<td>279, 280</td>
</tr>
<tr>
<td>Vascularization</td>
<td>315</td>
</tr>
<tr>
<td>VEGF. See Vascular endothelial growth factor (VEGF)</td>
<td></td>
</tr>
<tr>
<td>White blood cells. See Monocytes</td>
<td></td>
</tr>
<tr>
<td>Whole-genome expression profiling</td>
<td>249</td>
</tr>
<tr>
<td>Wnt/β-catenin</td>
<td>272</td>
</tr>
<tr>
<td>Wnt signaling</td>
<td>266, 273</td>
</tr>
<tr>
<td>World Wide Web for Cancer Research</td>
<td>317</td>
</tr>
<tr>
<td>Wound-healing, gene phenotype</td>
<td>200</td>
</tr>
<tr>
<td>WWOX expression</td>
<td>56</td>
</tr>
<tr>
<td>Xenograft model</td>
<td>265</td>
</tr>
<tr>
<td>Xenotransplantation</td>
<td>264, 269</td>
</tr>
<tr>
<td>X-ray mammography</td>
<td>297</td>
</tr>
<tr>
<td>Xylene</td>
<td>218</td>
</tr>
</tbody>
</table>
Pharmacogenetics is becoming increasingly relevant in the diagnosis, treatment, and recovery of cancer patients. A major problem facing oncologists is the outstanding varied efficacy of treatment. Promising advances in pharmacogenetics have allowed the development of effective agents which will enable personalized cancer chemotherapy to become routine for the clinical practice.

Written by experts in the field and combining information that is unable to be found in a single source volume, Pharamcogenetics of Breast Cancer:

- combines a complete overview of pharmacogenetics and how it relates to breast oncology for diagnosis, treatment, and the recovery of patients using an individualized therapy model
- is the first single source reference demonstrating the application of pharmacogenetics in the care of breast cancer patients
- enables physicians a coherent interpretation of the emerging science of pharmacogenetics, aiding them to incorporate individual therapies in their own practice
- gives practical guidance on various forms of specimen collection, tissue selection, and handling procedures

BRIAN LEYLAND-JONES is Associate Vice President and Director, Emory Winship Cancer Institute, Emory University, Atlanta, Georgia. Dr. Leyland-Jones’ main research interests are pharmacodynamics, pharmacokinetics, and pharmacogenetics in oncological clinical trials; translation of preclinical models into the clinic; biomarker endpoints in Phase I/II clinical trials; and screening and mechanistic studies of novel targeted and chemotherapeutic anti-cancer agents. He has authored more than 125 peer-reviewed articles and book contributions, 150 abstracts, and 29 patents.