Published in final edited form as:

Breast Cancer Res Treat. 2009 December ; 118(3): 635–643. doi:10.1007/s10549-008-0301-1.

Novel Sampling Strategies to Enable Microarray Gene Expression Signatures in Breast Cancer: A Study to Determine Feasibility and Reproducibility in the Context of Clinical Care

Christopher L. Tebbit, Jun Zhai, Brian R. Untch, Matthew J. Ellis, Holly K. Dressman, Rex C. Bentley, Jay A. Baker, Paul K. Marcom, Joseph R. Nevins, Jeffrey R. Marks, and John A. Olson Jr²

Institute of Genome Sciences and Policy, Duke University, Durham, NC [J.Z., A.B., H.D., J.R.N., J.A.O.]; Department of Surgery, Duke University Medical Center, Durham, NC [C.L.T., B.R.U., J.R.M., J.A.O.]; Department of Medicine, Washington University School of Medicine, St Louis, MO [M.J.E.]; Department of Pathology, Duke University Medical Center, Durham, NC [R.C.B.]; and Department of Radiology, Duke University Medical Center, Durham, NC [J.B.]

Abstract

Purpose—Feasibility and reproducibility of microarray biomarkers in clinical settings are doubted because of reliance on fresh frozen tissue. We sought to develop and test a paradigm of frozen tissue collection from early breast tumors to enable use of microarray in oncology practice.

Experimental Design—Frozen core needle biopsies (CNBx) were collected from 150 clinical stage I patients during image-guided diagnostic biopsy and/or surgery. Histology and tumor content from frozen cores were compared to diagnostic specimens. Twenty eight patients had microarray analysis to examine accuracy and reproducibility of predictive gene signatures developed for estrogen receptor (ER) and HER2.

Results—One hundred twenty seven (85%) of 150 patients had at least one frozen core containing cancer suitable for microarray analysis. Larger tumor size, *ex vivo* biopsy, and use of a new specimen device increased the likelihood of obtaining representative specimens. Sufficient quality RNA was obtained from 90% of tumor cores. Microarray signatures predictive ER and HER2 expression were developed in a training sets of up to 356 surgical samples and were applied to microarray data obtained from core samples collected in clinical settings. In these samples, a sensitivity / specificity of 94% / 100% and 82% / 72% for predicting ER and HER2, respectively was achieved. Predictions were reproducible in 83–100% of paired diagnostic and surgical samples.

Conclusions—Frozen CNBx can be readily obtained from most breast cancers without interfering with pathologic evaluation. Collection of tumor tissue at diagnostic biopsy and/or at surgery from lumpectomy specimens using image guidance resulted in sufficient samples for array analysis from over 90% of patients. Sampling of breast cancer for microarray data is reproducible and feasible in clinical settings and can yield signatures predictive of multiple breast cancer phenotypes.

²To whom requests for reprints should be addressed, Box 3873, Duke University Medical Center, Durham, NC 27710. Phone: (919) 684.6133; Fax: (919) 681.6291; jaomd@duke.edu.

Declaration of Potential Conflicts of Interest

Duke University has been issued a U.S. patent (7,172,558 B2) for the tissue sampling device.

Dr. Olson is founder of Core Prognostex, Inc. a company that sells tissue collection kits for correlative studies in oncology clinical trials. This company had no involvement in this study.

Introduction

Analysis of breast cancer microarray data can identify patterns of gene expression that subclassify tumors of similar histology and predict their clinical behavior (1–9). Our initial study established methods to distinguish ER positive from negative and LN positive from negative using groups of genes, termed metagenes. Subsequently, by combining metagene patterns with clinical data we were able to improve discrimination of highly LN positive (>10) from negative patients (10, 11). Most recently we have used microarray analysis to develop expression signatures characteristic of a variety of oncogenic pathways for the purpose of predicting sensitivity to pathway specific agents (12, 13). For each type of prediction, different sets of genes yielded the best discrimination highlighting the utility of interrogating the entire genome rather than a small subset of genes.

While methods have been developed to analyze expression of a small subset of genes in formalin fixed material by PCR (e.g., Oncotype Dx), generation of a full expression profile greatly benefits from high quality RNA associated with rapid specimen freezing after devascularization. The average size of incident breast cancers has decreased over the last three decades, largely attributed to increased screening and awareness. Further, the wide-scale adoption of breast conserving surgery has made the evaluation of surgical margins one of the most important pathologic criteria after cytoreduction. Both of these factors have made it more difficult to routinely obtain a rapidly frozen sample of the tumor for research or clinical assays. In order to apply genomic technologies as clinical tests a reliable method of high quality fresh frozen sample collection from most incident cancers is required.(14) The purpose of this study was to develop and test a clinically applicable method of sampling early stage breast tumors for microarray studies using core biopsy and demonstrate its utility in evaluating microarray models of breast cancer clinical outcomes. Our results show that a sampling strategy using CNBx, either at the time of routine image-guided diagnostic biopsy or at time of definitive surgery is an effective and reproducible method of obtaining frozen tissue for array analysis without compromising the pathologic assessment of the resected tumor. Further, such samples may be used to generate reproducible microarray profiles suitable for inclusion in clinicogenomic outcome models.

Methods

Subjects and tissue

Two breast cancer tissue banking protocols at Duke University Medical Center included women with clinical stage I or II primary breast cancer who had clinically or radiographically measurable primary tumor. The surgical protocol used between 1987 and 2001 relied upon banking of specimens in surgical samples following resection. The CNBx protocol was developed in 2001 and included patients undergoing image guided biopsy or surgical excision of their breast cancer. All patients gave written informed consent for participation in these IRB-approved studies.

The CNBx protocol allowed collection of 14-gauge core needle tumor specimens at radiological biopsy and/or surgical excision (Supplemental Figure 1). In all cases, research specimens were obtained only after acquisition of diagnostic specimens. At surgery, research tumor cores were taken from lumpectomy or mastectomy specimens by the operating surgeon *ex-vivo* using a spring-loaded 14-gauge core needle biopsy (Achieve Needle, Cardinal Health, McGraw Park, IL). Collected research core biopsies were rapidly frozen in Tissue-Tek® optimal cutting temperature (O.C.T.) compound (Sakura Finetek USA, Inc, Torrance, CA) on-site using dry ice and frozen sections were obtained for histologic evaluation. The protocol emphasized an embargo of research tissue specimens until histologic comparison was made between research core frozen section and the

diagnostic specimen by the case pathologist to ensure that no clinically important diagnostic information was uniquely in the research core (e.g. in situ carcinoma in diagnostic specimen and invasive cancer in the research core).

Tumor staging by AJCC criteria and clinical biomarker assignment from study cases was performed by institutional pathologists in laboratories certified under the Clinical Laboratory Improvements Amendments of 1988 (CLIA). Biomarker analysis of breast cancers was performed according to the standard clinical protocols as follows. Assignment of ER status was made on formalin-fixed, paraffin-embedded tissue sections using ER monoclonal antibodies from Signet Laboratories (clone 1D5) and an avidin-biotin complex detection method (BioGenex, San Ramon, CA). Quantitation of ER expression was determined using a CAS200 digital image analyzer. Using a standard curve prepared from tumor with known amounts of ER and PR, the image analyzer converts the receptor data into fmol/mg cytosolic protein. Determination of HER2 overexpression was made by immunohistochemistry using rabbit anti-human c-erbB-2 antibody (DAKO, Carpinteria, CA). Cases deemed 2+ (intermediate intensity of staining) by immunohistochemistry (IHC) were further evaluated using fluorescence in situ hybridization (FISH) analysis for HER2 gene amplification (Vysis PathVysion Kit, Downers Grove, IL).

For microarray gene expression analysis, total RNA was extracted from tumor samples using Qiagen (Valencia, CA) reagents. Microarray data was generated from extracted RNA with Affymetrix Hu95av2 Genechips according to the manufacturer's instructions (Affymetrix, Santa Clara, CA). For two samples an additional round of amplification was performed according to manufacturer's protocol (Ambion, Austin, Tx). Hybridization targets (probes for hybridization) were prepared from total RNA according to standard Affymetrix protocols. Scanned data was analyzed using the Affymetrix Microarray Analysis Suite (v5.0).

For ER and HER2 predictive model development we used a Shotgun Stochastic Search (SSS) algorithm as previously described (15). Primary tumor derived expression analysis focused on subsets of the 436 available patient surgical samples based on tumor types and the availability of relevant biomarkers and clinicopathologic data (Supplemental Table 1). Training models were evaluated using a leave one out cross validation approach as previously described (1). The most robust training model for each given phenotype (e.g., ER and HER2 status) was then applied to the independent set of core biopsy samples.

Statistical analyses of tissue collection and clinical variables were performed using JMP software (SAS, Cary, NC). Categorical variables were compared using the Fisher's Exact Test and continuous variables were compared using the Student's T-test.

Results

This study was initiated in the fourth quarter of 2001 to respond to two challenges: 1) Banking of frozen breast cancer specimens for research at Duke that began in 1987 had tapered off dramatically beginning in 1998 even though surgeries for breast cancer had increased (Figure 1). For the 12 years including 1998, we banked an average of 107 cancers per year (range, 80–159). In the subsequent three years from 1999 through 2001, this number dropped to 44 cases per year (range, 27–62) using the same banking protocols. 2) Expression microarray analysis promised insights into the biology, response, and outcome of breast cancer but to obtain meaningful results, rapidly frozen tissue would be required on most incident cancers.

As a feasibility test for the CNBx tissue acquisition procedure, 150 consecutive consented cases were analyzed for tumor collection outcomes (Table 1). Tumor stage was T0 in 11

patients, T1 in 85 patients (57%), T2 in 40 patients (27%), and either T3 or T4 in 14 patients (16%). Median tumor size was 1.6 cm (range 0.1–10 cm). Compared to the tissues obtained by the Duke Breast Tumor Bank using a surgical banking protocol from 1987–2001, tumors in the CNBx group were of significantly lower T stage (T1 57% vs. 37%, $p<0.001$ / T2 27% vs. 37%, $p=0.011$) with a smaller mean size, 1.6 versus 2.0 cm, ($p=NS$) (Figure 2). There was no difference in the percentage of ER or HER2 expressing cancers between the groups.

An average of 4.2 research core biopsies was obtained from each patient (range 1–17, Table 1). As indicated above, a stained section of each research core was analyzed by the study pathologist and compared to the diagnostic specimens. From this comparison, we found an overall concordance of 85% (127/150) between research and diagnostic tissues. Since different cores from the same cancer (both research and diagnostic) often contained varying degrees of histology (benign, DCIS, invasive), this number represents the concordance using each set of specimens as a group rather than considering them individually. We further compared whether cores obtained at the time of surgery were more likely to contain concordant pathology than those obtained at image-guided biopsy. While we did obtain significantly more cores at surgery (mean 4.4 vs 2.4, $p=0.001$), the degree of concordance was not significantly increased ($p=0.1$). For cancers sampled in both radiology and surgery, research cores containing the diagnostic pathology were nearly always obtained (27/28, $p<0.0001$).

For generating useful gene expression data via microarrays, tumor content is considered to be an important parameter. Breast cancer is markedly heterogeneous and in all cases presents a challenge in this area. In our previous studies, we have used 60% tumor content (on a per cell basis) as an arbitrary cutoff to accept a specimen for RNA extraction without micro or macro dissection (1,10,11). Tumor cell content for core specimens ranged between 0–80% on a per cell basis and the number of cores with >60% tumor content was greater for T2 than T1 tumors (mean, 2.0 vs. 1.2, $p=0.003$). Ninety-one (61%) of the 150 patients had at least one core containing 60% cancer while 127 (85%) had a core specimen containing any amount of cancer. For T1 tumors a core containing 60% cancer was obtained 58% (49/85) of the time and a core containing any percentage of invasive cancer was collected 82% (70/85) of the time. At surgical resection, more core samples were collected and more of these cores had greater than 60% cancer in them (1.5 vs. 0.7, $p=0.001$) than those collected at diagnostic biopsy.

During this feasibility test we realized that while there was no danger to the patient in coring a lumpectomy specimen *ex vivo*, there was a significant risk of injury to the person performing the procedure. Immobilization of the specimen by hand is difficult and does not provide the stability or guidance to accurately core a small tumor. To solve this problem, we designed and fabricated a device for image-guided *ex vivo* core biopsy of small breast tumors contained within excised lumpectomy specimens (Figure 3A). Using this device, breast lumpectomy specimens were immobilized in a defined orientation with little compression and imaged with a specimen radiogram to identify the location of the tumor within the specimen (Figure 3B). Spring-loaded core needle biopsies were immediately taken and snap frozen in OCT (Figure 3C). Use of this device yielded tumor containing cores in 17 of 18 (94%) patients as compared to 53 of 75 (71%) without the device ($p=0.037$). Use of the stereotactic device was also associated with a greater number of cores procured per specimen and a significantly increased ability to obtain cores containing at least 60% tumor ($p=0.001$) (Table 1). The acquisition of core needle samples using this method did not interfere with standard pathological procedures and margin integrity was preserved.

One of the primary goals of this work was to determine whether an approach could be developed to yield gene expression array data on the majority of incident breast cancers. Forty-one core biopsy samples with $\geq 60\%$ tumor content were selected from 28 patients with T1 or T2 tumors collected at radiologic biopsy, surgery, or both. Extraction of total RNA from the 41 core samples yielded from 2.2 to 29.0 $\mu\text{g}/\text{core}$ with an average of 10.4 μg (Table 2). RNA integrity was evaluated using the Agilent Bioanalyzer with 2 samples showing substantially degraded material. The remaining 39 RNA samples from 27 patients were used to prepare probe that was applied to Affymetrix Hu95Av2 arrays using standard methods. Scaling factors for these arrays were all within the acceptable range and the raw fluorescent intensities were converted to expression values using the MAS5 algorithm.

Numerous gene expression signatures in breast cancer have been published based on the Affymetrix platform, however none are in routine clinical use. To evaluate the ability for microarray data to predict breast cancer phenotypes we chose to model parameters that are currently measured clinically by other methods, estrogen receptor (ER) and HER2 status. Successful prediction of these biomarkers using single Affymetrix probes for ER and HER2 and predictive gene signatures has been reported by others.(16, 17) To further examine this concept and to examine reproducibility of predictive signatures, we chose to develop predictive models for these biomarkers using the metagene approaches previously published by our group. For ER, a total of twenty gene probes were employed in the model (Supplemental Table 2). Predictive genes include known estrogen responsive genes LIV-1, IL6ST and CCND1 (18–21). Using a leave one out cross validation (LOOCV) approach in a training set of 363 samples, the sensitivity was 0.88 and the specificity 0.81 using clinically assigned ER status as truth (Figure 4A). When applied to the test set of 27 patients, the sensitivity and specificity of the ER predictor was 0.94 and 1.0, respectively (Figure 4B, top panel). To better estimate clinical utility of this model we examined 90% confidence intervals around the predictions. Estrogen receptor status was confidently (i.e., confidence interval excluding 0.5) predicted in 24/27 (89%) instances while 3/15 ER positive patients had confidence intervals crossing the 0.5 probability threshold. Of the 27 samples, the predictive probability of true positive was greater than 0.9 in 6/15 ER positive samples and was less than 0.1 in 8/11 ER negative samples. Probability of ER expression by the model correlated positively (correlation coefficient of 0.83) with the Allred score applied to estrogen receptor staining (data not shown).

Next we examined the ability of gene expression models to predict HER2 status as assigned clinically by combined IHC and FISH. For HER2, a total of 9 gene probes were used in the final model including HER2, RGL2, RGS16, and CLI (supplemental Table 3) (22–25). Using LOOCV in a training set of 236 samples, a sensitivity of 0.73 and a specificity of 0.87 were achieved (Figure 5A). The set of 24 core biopsy samples were used after removing three patients with lobular histology (pts 10, 24, and 26). For the remaining samples, the sensitivity and specificity of the HER2 classifier were 0.82 and 0.72 respectively. A prediction of HER2 status was made confidently (90% intervals not including 0.5) in 14/24 patients (58%). Of those 10 samples confidently predicted negative, 8 were true negatives based on clinical assignment. Of 4 samples confidently predicted positive, all 4 were HER2 positive by IHC and/or FISH (Figure 5B, top panel). The predictive probability of true positive was greater than 0.9 in 4/5 HER2 positive samples and was less than 0.1 in 2/18 HER2 negative samples. Ten of 24 patient samples were predicted with confidence intervals crossing the 0.5 probability threshold.

To begin to assess biologic and technical reproducibility, we included a series of replicate specimens in this study. Among the 27 individual patients, 5 donated specimens at both radiologic biopsy and weeks later at the time of tumor resection. From a separate set of 5 patient specimens we examined the effect of an extended period (2 hrs) of room temperature

ischemia on gene expression and predictions compared to a parallel specimen that was frozen immediately upon devascularization (non-ischemic). This time point was chosen as typical of the processing interval that might be encountered in most tissue acquisition protocols. Further, two cases were used to compare the effects of a second round of probe amplification that would be required with the small amounts of RNA derived from microdissected specimens, e.g., samples with a sparse tumor content. Prediction of ER status was concordant with similar confidence intervals in 12/12 cases (Figure 4B, lower panel). For HER2, predictions were concordant in 10/12 cases (Figure 5B, lower panel). Discordant predictions occurred between one radiology/surgery pair (25R-S) and one ischemic/non-ischemic pair (11NI-I).

Discussion

The ability to utilize powerful genomic-scale methodologies such as microarrays to develop predictive tools that can address the heterogeneity of cancer phenotypes has been demonstrated in numerous studies (1–4). For these methodologies to become clinically useful, practical issues of feasibility of frozen specimen collection and reproducibility of expression data must be demonstrated. For breast cancer specifically, there are several logistical hurdles that need to be overcome before routine acquisition of frozen tissue for genome level diagnostics is practical. We considered that such a method must: 1) Integrate seamlessly with current clinical care and standard pathologic assessment of tumor specimens; 2) Be applicable to T1 tumors, the most commonly encountered breast cancers in the U.S.; 3) Minimize process- and biologic-related variability in tumor gene expression; and 4) Require relatively little additional effort from staff to implement.

The primary reason that the ability to bank frozen breast cancer specimens became difficult was the shift away from partial mastectomy and towards breast conserving surgery. From the standpoint of surgical pathology, beyond the basic histologic diagnosis, the key determinant from a lumpectomy specimen is whether the surgical margins are clear of cancer. This requires intact margins preventing the simple wedging or bisection of the specimen for banking. If the specimen is obtained late in the day or needs to be transported to the surgical pathology lab (as is often the case), the most expeditious and safest course of action to maintain the geometry of the specimen is to immerse it in formalin for subsequent gross and microscopic pathology.

To address this problem, we designed a protocol that allows collection of frozen specimens from most incident cancers without compromising patient safety, clinical workflow, or pathologic diagnosis. The key elements of this protocol are: 1) Obtaining additional cores at the time of image guided diagnostic biopsy whenever possible, 2) The use of a biopsy device to core lumpectomy/segmentectomy specimens immediately after tissue devascularization, 3) All cores obtained with this research protocol are sectioned, stained, and examined by the study pathologist. Additionally, the research cores are handled with the same chain of custody procedures used for clinical diagnostic specimens. These specimens are released for research purposes only after it is determined that they do not contain pathologic information not present in the diagnostic cores. For example, if the diagnostic cores demonstrate only in situ disease while the research core contains invasive cancer, the research cores must be included in the official diagnostic pathology work up. Even though the research cores would not have been obtained without this procedure, this “embargo” step ensures that all available information will be used for diagnostic purposes. To date, no patient has had a compromise of their clinical diagnosis by participation in this protocol.

Our results show that sufficient frozen tumor core biopsies can be collected in clinical settings to allow routine acquisition of microarray data for most patients with early stage

breast cancers. Factors associated with successful frozen tumor procurement included larger tumors, use of image guidance to collect cores, and collection of samples at both the time of diagnostic biopsy and surgery. Using our protocol, samples sufficient for generating microarray data were obtained from over 90% of patients with breast tumors. Nevertheless, a significant minority of cases yielded insufficient cancer tissue for microarray analysis indicating that additional steps including microdissection and a second round of probe amplification may be required for routine and broad clinical application of this technology, particularly from T1 tumors. This is particularly important when complementary protein-based assays are proposed from limiting tissue resources. Our data also show that 2 hours of tissue ischemia at room temperature does not significantly reduce the quality of RNA or the expression profile of the tumor. This is a significantly longer period of time than we permitted in the current protocol but may be more typical of processing times encountered in other settings. Further, the development of a tumor specimen collection device improved the ability to get quality specimens from smaller tumors.

A significant test of the performance of samples collected by these biopsy procedures is the capacity to yield gene expression data that has predictive value equivalent to that obtained from primary tumor samples from traditional surgical specimens. Our data show that microarray based prediction of ER is both accurate and reproducible while HER2 assessment was less accurate, but still reproducible. The advantage of the predictive models shown is that both probabilities of a phenotype are expressed (e.g. ER probability of positive 0.9) as well as the confidence interval around the prediction are shown. We feel this method is of great potential clinical utility by conveying both the strength and uncertainty of the prediction. Predictions in the paired biopsy specimens also performed consistently with concordance of ER and HER2 assignments ranging from 83–100%.

As microarray data begins to play a larger role in therapeutic trials the importance of sample collection and treatment consistency will be paramount. Our data indicate that these differences while notably minor in this study should be anticipated and accounted for in a given study design. It is unclear whether the rate of discordance that we observed in microarray data is significantly different than would be encountered in a similar analysis of single analytes, e.g., ER and HER2 status measured by IHC in two different specimens, where concordance rates of 80–97% are reported, depending upon assay. (26, 27) In general, our reproducibility data show that microarray profiles from core biopsies are similar to those obtained from resected specimens and provide reassurance that the “same answer” can be obtained from cancer specimens sampled at different times and in different conditions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by the James Ewing Oncology Fellowship Award for Basic Research from the Society of Surgical Oncology (JAO), NIH K23 CA106595 (JAO), R21 CA108707 (JAO), NIH T32 CA93245-01 (CLT), and the Duke/DFCC breast SPOREs.

The authors thank Mike West, PhD, for discussion of data normalization and statistical approaches to microarray data used in this project. We also acknowledge Andrea Richardson, MD, PhD, and J. Dirk Iglehart, MD for microarray data sharing as part of an NIH breast cancer inter-SPORE collaboration between Duke University and the Dana Farber Cancer Center.

Reference List

1. West M, Blanchette C, Dressman H, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci USA*. 2001; 98:11462–7. [PubMed: 11562467]
2. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*. 2002; 347:1999–2009. [PubMed: 12490681]
3. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002; 415:530–536.
4. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000; 406:747–52. [PubMed: 10963602]
5. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001; 98:10869–74. [PubMed: 11553815]
6. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med*. 2006; 355:560–9. [PubMed: 16899776]
7. Buchholz TA, Stivers DN, Stec J, et al. Global gene expression changes during neoadjuvant chemotherapy for human breast cancer. *Cancer J*. 2002; 8:461–8. [PubMed: 12500855]
8. Pusztai L, Ayers M, Stec J, et al. Gene expression profiles obtained from fine-needle aspirations of breast cancer reliably identify routine prognostic markers and reveal large-scale molecular differences between estrogen-negative and estrogen-positive tumors. *Clin Cancer Res*. 2003; 9:2406–15. [PubMed: 12855612]
9. Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res*. 2005; 11:5678–85. [PubMed: 16115903]
10. Huang E, Cheng SH, Dressman H, et al. Gene expression predictors of breast cancer outcomes. *Lancet*. 2003; 361:1590–6. [PubMed: 12747878]
11. Pittman J, Huang E, Dressman H, et al. Integrated modeling of clinical and gene expression information for personalized prediction of disease outcomes. *Proc Natl Acad Sci USA*. 2004; 101:8431–36. [PubMed: 15152076]
12. Huang E, Ishida S, Pittman J, et al. Gene expression phenotypic models that predict the activity of oncogenic pathways. *Nat Genet*. 2003; 34:226–30. [PubMed: 12754511]
13. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature*. 2006; 439:353–7. [PubMed: 16273092]
14. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer*. 2007; 7:545–53. [PubMed: 17585334]
15. Dressman HK, Berchuck A, Chan G, et al. An integrated genomic-based approach to personalized treatment of patients with advanced-stage ovarian cancer. *J Clin Oncol*. 2007; 25:517–25. [PubMed: 17290060]
16. Gong Y, Yan K, Lin F, et al. Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. *Lancet Oncol*. 2007; 8:203–11. [PubMed: 17329190]
17. Pusztai L, Ayers M, Stec J, et al. Gene expression profiles obtained from fine-needle aspirations of breast cancer reliably identify routine prognostic markers and reveal large-scale molecular differences between estrogen-negative and estrogen-positive tumors. *Clin Cancer Res*. 2003; 9:2406–15. [PubMed: 12855612]
18. Manning DL, McClelland RA, Knowlden JM, et al. Differential expression of oestrogen regulated genes in breast cancer. *Acta Oncol*. 1995; 34:641–6. [PubMed: 7546832]
19. Manning DL, McClelland RA, Gee JM, et al. The role of four oestrogen-responsive genes, pLIV1, pS2, pSYD3 and pSYD8, in predicting responsiveness to endocrine therapy in primary breast cancer. *Eur J Cancer*. 1993; 29:1462–8. [PubMed: 8398276]
20. Deb S, Tessier C, Prigent-Tessier A, et al. The expression of interleukin-6 (IL-6), IL-6 receptor, and gp130-kilodalton glycoprotein in the rat decidua and a decidual cell line: regulation by 17beta-estradiol and prolactin. *Endocrinology*. 1999; 140:4442–50. [PubMed: 10499497]

21. Altucci L, Addeo R, Cicatiello L, et al. 17beta-Estradiol induces cyclin D1 gene transcription, p36D1-p34cdk4 complex activation and p105Rb phosphorylation during mitogenic stimulation of G(1)-arrested human breast cancer cells. *Oncogene*. 1996; 12:2315–24. [PubMed: 8649771]
22. Derrien A, Druet KM. RGS16 function is regulated by epidermal growth factor receptor-mediated tyrosine phosphorylation. *J Biol Chem*. 2001; 276:48532–8. [PubMed: 11602604]
23. Isomura M, Okui K, Fujiwara T, Shin S, Nakamura Y. Isolation and mapping of RAB2L, a human cDNA that encodes a protein homologous to RalGDS. *Cytogenet Cell Genet*. 1996; 74:263–5. [PubMed: 8976381]
24. Xia W, Chen JS, Zhou X, et al. Phosphorylation/cytoplasmic localization of p21Cip1/WAF1 is associated with HER2/neu overexpression and provides a novel combination predictor for poor prognosis in breast cancer patients. *Clin Cancer Res*. 2004; 10:3815–24. [PubMed: 15173090]
25. Zhang H, Kim JK, Edwards CA, Xu Z, Taichman R, Wang CY. Clusterin inhibits apoptosis by interacting with activated Bax. *Nat Cell Biol*. 2005; 7:909–15. [PubMed: 16113678]
26. Wood B, Junckerstorff R, Sterrett G, Frost F, Harvey J, Robbins P. A comparison of immunohistochemical staining for oestrogen receptor, progesterone receptor and HER-2 in breast core biopsies and subsequent excisions. *Pathology*. 2007; 39:391–5. [PubMed: 17676479]
27. Mann GB, Fahey VD, Feleppa F, Buchanan MR. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. *J Clin Oncol*. 2005; 23:5148–54. [PubMed: 16051956]

Future Application of the Work to Clinical Cancer Medicine

Numerous studies have shown the potential for microarray-based biomarkers to identify molecular subtypes of breast cancer with clinical implications. Most studies have relied on specimens retrieved from existing tumor banks in which samples (often from large tumors) were collected under varying conditions. For microarray diagnostics to become feasible in breast cancer care, issues of reliable and controlled fresh tumor collection from commonly encountered (often small) tumors must be addressed. Further, microarray profiles must be shown to be reproducible under a variety of tumor sampling conditions. We have described and tested a robust, clinically applicable method of tumor sampling. These results are now being applied to prospective validation studies of microarray biomarkers at single institution and cooperative group studies and will eventually allow implementation of fresh tissue based biomarkers to the future practice of cancer medicine.

Figure 1.

Breast cancer frozen specimen acquisition at Duke University Medical Center between 1987 and 2003. Upper panel shows total cases banked and total breast surgeries performed over study period. Lower panel shows percentage of cases banked during the same time period. The number of breast cancer surgeries was obtained from the Duke Comprehensive Cancer Center Tumor Registry.

Figure 2.

Stage comparison of banked tumor specimens before (dark columns) and after (open columns) development and implementation of the CNBx protocol. Data are expressed as percentage of patients with a given AJCC tumor stage in each group.

Figure 3. Stereotactic biopsy device for ex vivo collection of CNBx specimens from lumpectomy specimens. A. Lumpectomy specimen is fixed in a defined orientation within the device. B. A specimen radiogram is taken to identify the tumor within the lumpectomy specimen (yellow arrow) and orient it according to grid lines and coordinates on the device. C. 1 gauge CNBx are taken from the tumor through side apertures in the device.

Figure 4.

Microarray predictions of breast cancer ER status. A 20 gene classifier (supplemental table 2) was developed and evaluated using leave one out cross validation in a training set of 436 surgical breast cancer samples (A). Blue and red closed circles indicate predicted probabilities of ER negative and positive samples, respectively. The 90% confidence intervals of predictions are indicated with error bars. The ER model was then applied to an independent validation set of 27 CNBx breast cancer samples (B, top panel). Predictions of ER status in paired CNBx samples (radiology-surgery, amplified – unamplified, and ischemia-non-ischemia) are shown (B, lower panel). Blue and red open circles indicate predicted probabilities of ER negative and positive samples, respectively. The 90% confidence intervals of predictions are indicated with error bars.

A**Training set (N=236)**Sensitivity 73%
Specificity 87% **B****Test set (N=24)**Sensitivity 82%
Specificity 72% **Figure 5.**

Microarray predictions of breast cancer HER2 expression. A 9 gene classifier (supplemental table 3) was developed and evaluated using leave one out cross validation in a training set of 236 surgical breast cancer samples (A). Blue and red closed circles indicate predicted probabilities of ER negative and positive samples, respectively. The 90% confidence intervals of predictions are indicated with error bars. The Her2 model was then applied to an independent set of 27 CNBx breast cancer samples (B, top panel). Predictions of Her2 status in paired biopsy samples (radiology-surgery, amplified – unamplified, and ischemia-non-ischemia) are shown (B, lower panel). Blue and red open circles indicate predicted probabilities of ER negative and positive samples, respectively. Green circle indicates sample with borderline Her2 expression by FISH. The 90% confidence intervals of predictions are indicated with error bars.

Table 1

Variables associated with successful collection of frozen breast cancer CNBx for microarray analysis.

Variable	N	Avg# core s/pt	Avg# cores w/ 60% cancer	# pts with a 60% cancer core	# pts with any % cancer in core
T Stage					
T0 (DCIS)	11	3.5	0.0	1	5
T1	85	3.8	1.2	49	70
T2	40	5.1	2.0	31	40
T3	8	5.0	2.3	6	7
T4	6	3.0	1.0	4	5
Procurement site					
Radiology	57	2.4	0.7	29	45
Surgery	65	4.4	1.5	42	55
Both	28	7.1	2.7	20	27
Biopsy Device					
Yes	18	6.2	2.3	15	17
No	75	4.9	1.7	19	53
Total	150	4.2	1.4	91	127

* vs. T1 tumors;

** vs. radiology-acquired specimen;

vs. device not used.

Table 2

Yield of RNA from breast cancer CNBx.

Variable	Pts	Cores	Average RNA yield (ug)	Cores yielding 3 ug RNA	Pts with > 3 ug RNA
T stage					
T1a/T1s	1	4	10.5	1	1
T1b	2	2		2	2
T1c	13	12		12	13
T2	9	17	10.5	16	9
T3	3	6	8.2	5	3
Core tumor content					
22-50%		6	8.0	6	
51-80%		28	9.4	21	
> 80%		7	15.0	6	
Total	28	41	10.4	34	28