Evaluation of a Panel of Tumor Markers for Molecular Detection of Circulating Cancer Cells in Women with Suspected Breast Cancer

Monica M. Reinholz,1 Andrea Nibbe,2 Leslie M. Jonart,1 Kathleen Kitzmann,1 Vera J. Suman,2 James N. Ingle,3 Raymond Houghton,4 Barbara Zehentner,5 Patrick C. Roche,1 and Wilma L. Lingle1

Abstract  Purpose: We examined the feasibility of using molecular characterization of circulating tumor cells as a method for early detection of breast cancer.

Research Design: Women without a prior history of cancer who had a breast abnormality detected on imaging followed by a breast biopsy were enrolled in this study. Density gradient centrifugation and immunomagnetic capture were used to enrich for epithelial cells from ~20 mL of blood. Real-time reverse transcription-PCR was used to quantitate the expression levels of the highly breast-specific genes, mammaglobin, γ-aminobutyric acid type A receptor α subunit (GABA A α), B305D-C, and B726P in the epithelial cell-enriched samples.

Results: The assay was technically feasible in 154 of 199 accrued patients. From their clinical assessment, 100 patients had benign breast disease, 10 patients had ductal carcinoma in situ, and 44 patients had invasive breast cancer. We constructed a diagnostic test that classified patients with mammaglobin levels of at least 32.2 copies/pg γ-actin (units) in their circulating epithelial cells as positive for invasive breast cancer. This resulted in a sensitivity and specificity of 63.3% and 75.0%, respectively. A diagnostic test that classified patients as positive for invasive breast cancer when either mammaglobin levels were >46.3 units or B305D-C levels were >11.6 units increased the sensitivity and specificity to 70.5% and 81.0%, respectively. In the latter test, 12 of the 14 node-positive breast cancer patients were correctly identified. Including GABA A α and B726P in the test did not increase its diagnostic potential.

Conclusions: These results suggest that molecular characterization of circulating epithelial cells using mammaglobin and B305D-C offers potential for early detection of invasive breast cancer.

Early detection of breast cancer can reduce the morbidity and mortality of breast cancer patients (1–3). Breast cancer mortality occurs in a substantial portion of breast cancer patients (3) and is increasingly linked to early metastases, which are often occult at the time of primary diagnosis (4). Because undetected micrometastases can contribute to the failure of primary treatment (4), their identification in patients with early-stage breast cancer may have a substantial impact on prognosis and treatment choice for these patients. Thus, improved direct identification of these occult metastases in blood offers a critical opportunity to optimize management of breast cancer patients.

In addition to physical examination and mammography, sensitive molecular techniques may be used to detect early-stage breast cancer. Assays for nucleic acid–based markers are valuable tools for the sensitive detection and assessment of disease status in asymptomatic cancer patients (5, 6). Several techniques have been developed to enrich certain types of cells from blood and to characterize these cells using cytologic and reverse transcriptase-PCR (RT-PCR) assays (7). Application of these methods may allow for the early detection of cancer when the tumor burden is small and the disease is potentially more curable (5). It has been reported that anti-breast cancer epithelium monoclonal antibodies detected circulating cells in 95% of breast cancer patients before surgery and in 32% of the patients after surgery (8, 9). These promising results support the design of feasibility studies that will utilize sensitive molecular techniques to detect circulating breast cancer cells.

An immunomagnetic cell capture technique coupled with quantitative RT-PCR (qRT-PCR) has been developed for detection of breast cancer cells in blood and bone marrow (10). Using a test system with the MDA-MB-361 breast cancer cell line, we previously demonstrated that this method is extremely sensitive and is capable of detecting one breast cancer cell in 107 nucleated cells, which is equivalent to four cells per 10 mL blood (10). We hypothesize that sensitive molecular detection of cancer cells in peripheral blood using highly breast-specific genes (11–13) will provide a screening test that can be used independently or in...
concert with mammography and physical examination to accurately detect early-stage breast cancer.

In this study, we examined the blood of women with suspicious mammograms for evidence of circulating tumor cells using immunomagnetic cell capture and molecular detection for three known genes, cytokera tin-19 (CK-19), mammaglobin, and γ-am inobut yric acid type A receptor π subunit (GABA A), and two novel genes, B305D-C and B726P. The two latter genes were identified by differential display and subtractive hybridization (12). We examined CK-19 gene expression because CK-19 has been studied extensively as a potential marker for minimal residual disease in blood (14–25), bone marrow (24, 26–28), and lymph nodes (20, 21, 28–30). Although the functions of mammaglobin, B305D-C, B726P, and GABA A are not completely characterized, at least one of these genes was expressed in all 46 breast tumors and 50 metastatic breast cancer lymph node specimens examined (12). In addition, at least one of these genes was expressed in circulating epithelial cells isolated from 75% of 32 metastatic breast cancer patients (12). Notably, mammaglobin was detected in 62% of these patients.

In the current study, we compared our molecular detection results to the breast tissue biopsies to determine the concordance between the gene expression levels of CK-19, mammaglobin, B305D-C, B726P, and GABA A in circulating breast cancer cells and histologic findings. Our results showed that the combination of mammaglobin and B305D-C correctly identified 71% of the patients with invasive breast cancer, and, in particular, 86% of the node-positive (N+) breast cancer patients. Thus, detection of mammaglobin and B305D-C may constitute a sensitive, dual-marker screening assay for early detection of breast cancer. An additional study with an independent population of patients is necessary to validate this method as a beneficial screening test for breast cancer.

Materials and Methods

Patient specimens. Women with a breast abnormality detected on imaging and who were to undergo a breast biopsy were approached for this study. All patients were informed of the investigational aspects of this study and provided written consent in accordance with institutional and federal guidelines. One hundred ninety-nine patients with suspicious mammograms and subsequent tissue biopsies were enrolled in this study between July 16, 2001, and February 28, 2003. Approximately 20 mL of blood were collected from each patient and all blood samples were assigned a unique identification number. Board-certified anatomic pathologists at Mayo Clinic, Rochester, reviewed the original pathology findings. All assays were done with the investigators blinded to the results of the biopsy. Twenty patients were declared ineligible [5 had a prior history of invasive cancer, 14 had no breast tissue biopsy done (e.g., fine needle aspiration or ultrasound were done), and 1 had the blood draw after the biopsy]. One patient was also excluded because she had a recurrent phylloides-type tumor.

Of the 178 eligible patients, 24 (13%) patients were excluded due to technical difficulties in processing the blood samples. Six patients [four with benign breast disease, one with ductal carcinoma in situ (DCIS), and one with invasive breast cancer] were excluded because the β-actin mRNA levels were below 0.1 pg. Sixteen patients [14 with benign breast disease, one with DCIS, and one with invasive breast cancer] were excluded because of poor separation of the lysed epithelial cell suspension from the Epithelial Enrich Dynal beads (Dynal A.S., Oslo, Norway). Two patients [one with benign breast disease and one with invasive breast cancer] were excluded because of improper handling and storage of the samples. Thus, this study cohort consisted of 100 patients with benign breast disease (e.g., microcalcifications, fibrocystic and/or nonfibrocystic changes, and fibroadenomas), 10 patients with DCIS, and 44 patients with invasive breast cancer. The median age at biopsy was 51 years (range: 21-84 years) for those women with benign breast disease, 62 years (range: 41-75 years) for those women with DCIS, and 56 years (range: 37-81 years) for those women with invasive breast cancer.

Immunomagnetic enrichment of epithelial cells. Mononuclear cells were isolated from the blood sample and the epithelial cells were subseq uently captured from the mononuclear cell co el in the blood collection. The epithelial cells were enriched by immunomagnetic capture using the monoclonal antibody, Ber-EP4, and the magnetic Dynabeads Epithelial Enrich kit according to the manufacturer’s instructions (Dynal). The manufacturers showed that up to a 5 log enrichment of epithelial cells and a yield of 70% viable, bead-free tumor cells can be obtained using this kit (Dynal). The Ber-EP4 antibody recognizes two glycoproteins on the surface and in the cytoplasm of epithelial cells except the superficial layers of squamous epithelia, hepatocytes, and parietal cells (31).

Peripheral blood (18 mL) was added to an Accuspin Histopaque-1077 system (Sigma Aldrich, St. Louis, MO) and centrifuged at 1,500 rpm for 10 minutes in a Beckman CS-6R tabletop centrifuge (Beckman Instruments, Palo Alto, CA). The mononuclear cell layer was removed, washed twice with PBS, diluted to 1 mL with PBS/0.1% bovine serum albumin, and incubated with Epithelial Enrich Dynabeads (1 × 10^7 beads in a volume of 20 μL) while rocking for 1 hour. The suspension was placed on a magnet for at least 6 minutes and the supernatant was carefully removed. The cells attached to the magnetic beads were washed thrice with 1 mL PBS/0.1% bovine serum albumin and lysed with the lysis binding buffer supplied with the kit. The lysed cell suspension (with beads attached) was stored at −80°C until processing.

Messenger RNA isolation, reverse transcription, and second-strand synthesis. Messenger RNA was isolated from the cell lysate using the mRNA Direct kit and Dynabeads Oligo (dT)_25 according to the manufacturer’s instructions (Fig. 1A; Dynal). The lysed cell suspension was separated from the Epithelial Dynabeads and incubated with 10 μL of prewashed Dynabeads Oligo (dT)_25 (1.0 μg/10 μL) for 5 minutes while rotating on a mixer to ensure annealing between the lysate and the beads. This solution was placed on a magnet to remove the supernatant. The beads and the attached mRNA of each sample were washed multiple times in the appropriate buffers included with the kit.

The reverse transcription reaction was done in 20 μL containing the mRNA attached to the Dynabeads Oligo (dT)_25, 1× reverse transcriptase buffer, 1 mmol/L deoxynucleotide triphosphates, 25 units RNase inhibitor, and 40 units avian myeloblastosis virus reverse transcriptase (Roche Diagnostics, Indianapolis, IN). This suspension was gently vortexed, incubated at 42°C for 60 minutes (with gentle vortexing every 10 minutes), and held at 4°C. After the first-strand synthesis was completed, the resultant mRNA/cDNA bead complex was heated at 95°C for 1 minute in 2 mmol/L EDTA [5 μL of 0.01 mol/L EDTA (pH 7.5) was added to the 20 μL bead suspension] and the melted mRNA was immediately removed and discarded. Another heat treatment was done by the addition of 20 μL of 2 mmol/L EDTA (pH 7.5) and heated at 95°C for 1 minute, and any remaining mRNA was immediately removed and discarded. This process yielded first-strand cDNA covalently attached to the magnetic beads (Fig. 1B).

The first-strand cDNA bead complex was washed thrice in 50 μL of a buffer containing final concentrations of 5 mmol/L KCl, 1 mmol/L Tris (pH 8.3), 0.15 mmol/L MgCl2, and 0.1 mg/mL bovine serum albumin (Pierce, Rockford, IL) and stored in this same buffer at 4°C (32, 33).

The resultant solid-phase cDNA library was used repeatedly with gene-specific, forward primers to CK-19, mammaglobin, B305D-C, B726P, GABA A, and β-actin to synthesize gene-specific, second-strand cDNA (32, 33), which served as a template for real-time PCR (Fig. 1C; refs. 34, 35). The second-strand reaction was done in 20 μL.
containing the cDNA attached to the beads. PCR reaction buffer containing MgCl₂ (Roche Diagnostics), 0.1 mg/mL bovine serum albumin, 0.5 pmol/L of the specific forward primer, 0.2 μmol/L deoxynucleotide triphosphates, and 1 unit Taq polymerase (Roche Diagnostics). This suspension was slightly vortexed and one PCR cycle was done at 94°C for 20 seconds, 60°C for 2 minutes, 72°C for 10 minutes, and then held at 4°C. The bead suspension was incubated at 94°C for 2 minutes to melt off the gene-specific, second-strand cDNA, which was immediately removed and stored at −20°C until needed for real-time PCR.
Appropriate assay controls included a positive and a negative reverse transcriptase control using mRNA isolated from MDA-MB-361 breast cancer cells, which contain all genes of interest. Blood from normal healthy female donors was processed as described above and included in each assay run.

**Real-time PCR (quantitative PCR).** The second-strand cDNA from each sample was amplified with 40 cycles by fluorescent-based kinetic PCR with gene-specific primers using an ABI 7700 Sequence Detection System (PE Biosystems, Foster City, CA). Circulating epithelial cells were detected using mammaglobin, B305D-C, B726P, GABA Aα, and CK-19.

The sequences for the primers and fluorescent probes for B305D-C, B726P, and GABA Aα were obtained from Houghton et al. (12). Primers and fluorescent probes for mammaglobin and CK-19 were based upon previously published human sequence data (11, 36) and were designed using the Primer Express Program (Applied Biosystems, Foster City, CA). The primers and probes for all the genes except β-actin were purchased from Integrated DNA Technologies (Coralville, IA). The β-actin primers and probe were obtained from a kit purchased from Applied Biosystems. The sequences and positions of the nucleotide sets for each gene are shown in Table 1.

The PCR reaction was done in 25 μL containing the unknown sample template, primers and probes, and universal Master Mix (Applied Biosystems), which consisted of dATP, dCTP, dGTP, dUTP, Amplitaq Gold, Amperase UNG, and trace amounts of glycerol, Tween 20, and glycine. Each forward and reverse primer for mammaglobin, CK-19, and β-actin was used at a concentration of 300 nmol/L. The final probe concentration for these reactions was 200 nmol/L. The forward primer concentrations for B305D-C, B726P, and GABA Aα were 900, 300, and 900 nmol/L, respectively. The reverse primer concentration for each of these reactions was 900 nmol/L. The final probe concentration for these reactions was 160 nmol/L. The unknown sample templates were either 5 μL of the respective second-strand template, 5 μL of the respective plasmid, or 5 μL of water (as a no template control). Each reaction was done in triplicate except for the no template controls, which were done in quadruplicate. The universal thermal cycling program in the real-time PCR reaction consisted of an initial 2-minute incubation at 50°C to activate the UNG enzyme, a 10-minute incubation at 95°C to activate the AmpliTaq Gold, and 40 cycles of 95°C for 0.15 minutes and 60°C for 1 minute.

Plasmids containing the amplicon of interest for B305D-C, B726P, and GABA Aα were obtained from Houghton and coworkers, and the plasmid containing the β-actin amplicon was constructed using the Original TA Cloning kit according to the manufacturer’s instructions (Invitrogen, Carlsbad, CA). These plasmids or oligonucleotides of the specific amplicons (mammaglobin and CK-19) were used in the real-time PCR assays to generate standard curves to determine the level of gene expression. Standard curves were constructed using the Ct values (defined as the fractional cycle number at which the fluorescence generated by cleavage of the probe passes a fixed threshold above baseline) determined in the real-time PCR and a known amount of the plasmid containing the amplicon of interest or the oligonucleotide. The initial cDNA concentration in the assay (related to the amount of target message) of mammaglobin, B305D-C, B726P, GABA Aα, and CK-19 in the unknown samples was determined using these standard curves. Standard dilutions ranging from 1 to 1 × 10^8 copies of each plasmid or oligonucleotide were used for this purpose.

The presence of β-actin mRNA (standard range: 1-1 × 10^5 copies; 0.312 fg-31.2 pg) was used as an endogenous control in the samples to normalize the gene expression of mammaglobin, CK-19, B305D-C, B726P, and GABA Aα. No signal indicative of the presence of contaminating genomic retrospseudogenes of β-actin was detected in negative controls without reverse transcriptase (37). We previously used β-actin as a normalization factor, plasmids containing the specific amplicon of interest for standards, and Taqman chemistries in qRT-PCR to quantify the gene expression of several transforming growth factor β family members as well as the transforming growth factor β – inducible early gene and some of its target genes in breast cancer disease progression (38, 39).

Samples in the current study were considered not detectable if at least two of the triplicate Cts were 40 and the remaining replicate Ct was above 37. A sample was reanalyzed (gene-specific, second strand was regenerated and qPCR was done again in triplicate) if two of the triplicate Cts were 40 and the remaining Ct was below 37. These samples were considered detectable if the Ct of at least one of the run triplicates was below 37. The corresponding copy numbers from the original and rerun assays were averaged and used for data analyses.

We previously showed that immunomagnetic capture (using B9E4P antibody-coated magnetic beads, Dynal) combined with real-time RT-PCR (using the Light Cycler, Roche Diagnostics) is an extremely sensitive method and is capable of detecting one epithelial cell in 1^10^ nucleated cells in a model test system using MDA-MB-361 breast cancer cells (10). We now modified this method to include generation of a solid-phase cDNA library that can be repeatedly used to synthesize gene-specific, second-strand cDNA (32, 33) and qPCR using Taqman chemistries (34, 35) in the ABI 7700 Sequence Detection System. In a preliminary study, we examined the sensitivity and specificity of this modified method by performing a cell titration assay using MDA-MB-361 breast cancer cells. A quantity of 0, 1, 10, 100, and 1,000 MDA-MB-361 breast cancer cells per milliliter of blood were added to each 10 mL of blood obtained from healthy female donors. Mononuclear cells were obtained, the epithelial cells were enriched by immunomagnetic capture, mRNA was isolated and reverse transcribed, and gene-specific, second-strand cDNA was synthesized as described above. Appropriate assay controls included 10 ng MDA-MB-361 mRNA with avian myeloblastosis virus reverse transcriptase (positive) and 10 ng MDA-MB-361 mRNA without avian myeloblastosis virus reverse transcriptase (negative, no reverse transcriptase). These controls were processed along with the nonspiked and spiked blood samples as described above. Real-time RT-PCR (50 cycles) was done for mammaglobin (Fig. 2; cycles 16-40), B305D-C, B726P, GABA Aα, and CK-19 (data not shown). The PCR reactions for all samples were done in duplicate. The no template controls were done in quadruplicate. As illustrated in Fig. 2, mammaglobin gene expression in the 1 cell/mL blood sample was detected at an average Ct of 27.83, demonstrating that this method is extremely sensitive. This method is also very specific as the ΔΔCt, for the negative controls (0 cell, no reverse transcriptase, and no template control) remained below the threshold (maximum ΔΔCt was ~1.04 × 10^-2; data not shown), indicating that a nonspecific increase in fluorescence was not observed. Overall, the combination of immunomagnetic capture and qRT-PCR is a very sensitive and specific method, which can be used to determine the feasibility of using molecular characterization of circulating tumor cells for early detection of breast cancer.

**Statistical analyses.** Wilcoxon rank sum tests were used to test for differences in the distributions of gene expressions between the benign and invasive breast cancer biopsies. Associations between gene expression levels within a disease group were explored using Spearman rank correlation coefficients. A P value < 0.01 was considered significant.

Sensitivity and specificity analyses were used to assess the discriminating ability of mammaglobin, B305D-C, B726P, GABA Aα, and CK-19 alone and in combination. We examined the role of these five genetic markers in discriminating among women with benign breast disease and those with invasive breast cancer. A sample was said to be positive for invasive breast cancer if its gene expression value was above an experimental cut point. Receiver operator characteristic curves were constructed to visualize the effect of altering the experimental cut point on sensitivity and specificity. The experimental cut point that maximized both sensitivity and specificity was chosen to be the final cut point for the gene.

In an attempt to increase our ability to discriminate between benign breast disease and invasive breast cancer, we examined combinations of two genes. In this case, a sample was classified as positive for invasive breast cancer if the gene expression level of either of the two genes exceeded its specified cut point. For each combination of genes,
sensitivity and specificity were calculated at each possible pairing of experimental cut points. The pairing of cut points that maximized both sensitivity and specificity was selected.

Combinations of three or more genes were also examined. Due to the computational intensity of examining all possible combinations of the five genes using a grid search, classification trees were used to find promising combinations. A grid search was conducted on only these resulting combinations.

Sensitivity and specificity values resulting from predictions based on the same data used to fit a model are prone to being overly optimistic. To quantify this optimism, estimates of the sensitivity and specificity to be expected in data from independent trials were calculated using the bootstrapping method suggested by Efron and Tibshirani (40). These estimates were computed for the most promising combination of genes using 1,000 bootstrap samples.

Results

Patient characteristics. The nodal and menopausal status of the 154 patients are listed in Table 2 and the clinicopathologic characteristics of the 44 primary, invasive breast tumors are presented in Table 3. The majority of the invasive breast cancer patients were postmenopausal (66%) diagnosed with hormone receptor–positive (75%), node-negative (66%), invasive ductal carcinomas (86%) that were ≤2 cm (61%).

Gene expression distributions. The distribution, by biopsy result, of mammaglobin, B305D-C, B726P, GABA A$\alpha$, and CK-19 expression levels in the circulating epithelial cells are displayed in Fig. 3. The median CK-19, mammaglobin, and B305D-C mRNA levels were significantly higher in the patients with invasive breast cancer than those with benign breast disease ($P = 0.004, P < 0.0001$, and $P = 0.006$, respectively). The median CK-19 gene expression was ~3-fold higher in tumor cells from patients with invasive breast cancer (75.1 units) than that in patients with benign breast disease (28.1 units). The median mammaglobin gene expression was ~8-fold higher in circulating cells from patients with invasive breast cancer (53.0 units) than those with benign breast disease (6.6 units). The median B305D-C gene expression was ~4-fold higher in cells from patients with invasive breast cancer (2.4 units) than those with benign breast disease (2.0 units). No significant differences in the median transcript levels of B726P and GABA A$\alpha$ were observed between patients with benign breast disease biopsies and those with invasive breast cancer biopsies.

Both CK-19 ($P = 0.057$) and mammaglobin ($P = 0.022$) mRNA levels seemed to be lower in DCIS patients than patients with invasive breast cancer. Due to the limited number of DCIS patients and little power, these results need to be interpreted with caution.

Correlation analyses. Among the patients with benign breast disease biopsies, the gene expression levels were positively pairwise correlated with each other. The Spearman rank correlation coefficients ($r_s$) ranged from 0.26 (CK-19/mammaglobin; $P = 0.007$) to 0.61 (mammaglobin/B726P; $P < 0.0001$). B305D-C was positively correlated with mammaglobin ($r_s = 0.39; P < 0.0001$) and CK-19 ($r_s = 0.56; P < 0.0001$).

Among the patients with invasive breast cancer biopsies, neither CK-19 and B726P nor CK-19 and GABA A$\alpha$ were found to be significantly correlated. All remaining pairwise correlations were significant. These correlations ranged from 0.39 (mammaglobin/GABA A$\alpha$; $P = 0.009$) to 0.68 (GABA A$\alpha$/B726P; $P < 0.0001$). Notably, mammaglobin was positively correlated with CK-19 ($r_s = 0.65; P < 0.0001$).

Interestingly, mammaglobin was strongly correlated with both B726P ($r_s = 0.84; P = 0.003$) and GABA A$\alpha$ ($r_s = 0.97; P < 0.0001$) in patients with DCIS. Again, these results need to be interpreted with caution because of the limited number of DCIS patients.

Individual discriminatory ability of mammaglobin, B305D-C, B726P, GABA A$\alpha$, and CK-19. The ability of mammaglobin, B305D-C, B726P, GABA A$\alpha$, and CK-19 to discriminate between patients with benign breast disease and invasive breast cancer was examined using receiver operator characteristic curves. We then determined the expression level of each biomarker that maximized both sensitivity and specificity. At a cut point of 32.2 units, mammaglobin provided the best discrimination with sensitivity of 63.6% (95% confidence interval, 48.9-76.2) and specificity of 75.0% (95% confidence interval, 65.7-82.5). With this cut point, 71.5% of the patients were correctly identified (Table 4), including 86% of the N+ patients. A sensitivity and specificity of ~60% was associated with the gene expression levels of B305D-C, B726P, GABA A$\alpha$, and CK-19.

Discriminatory ability of the combination of mammaglobin and B305D-C, B726P, GABA A$\alpha$, or CK-19. Combinations of genes were examined to determine if a set of genes resulted in improved discrimination between benign breast disease and invasive breast cancer biopsies. Mammaglobin, at a level of 46.3 units, with B305D-C, at a level of 11.9 units, provided the best pairing with a sensitivity of 70.5% (95% confidence interval, 55.8-81.8%) and a specificity of 81.0% (95% confidence interval, 72.2-87.5%; Table 5). Over three fourths of the patients were correctly classified as either having benign breast disease or invasive breast cancer (Fig. 4) including 86% of the N+ patients. Other gene combinations, including mammaglobin with B726P and mammaglobin with CK-19, provided sensitivities between 57% and 69% and specificities between 75% and 85% (Table 5). Although the combinations of B726P

### Table 1. Taqman primer and probe sequences

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer</th>
<th>Position</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-actin</td>
<td>TCACCCACACTG TGCCCCATCTACGA</td>
<td>2,141</td>
<td>CAGCGGAACGC TCATTGCCAATGG</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>AGAAGTGGCAAGG ATGGTGAGAA</td>
<td>367</td>
<td>ACATGATAGCAGT TTTACAAATGT</td>
</tr>
<tr>
<td>CK-19</td>
<td>CACGGGCTGTAT GGAATT</td>
<td>1,121</td>
<td>CTTGAGGCAACAGAA ATTTTG</td>
</tr>
<tr>
<td>B305D-C form</td>
<td>AAAGCACAGAGT GGTGGAGGT</td>
<td>117,906</td>
<td>CCTAGACACAAATGG CTTCTC</td>
</tr>
<tr>
<td>B726P</td>
<td>TCTGGTTTCTCAT TTATTCATTATT</td>
<td>235</td>
<td>TGCCAGGAGGGGATTATCT</td>
</tr>
<tr>
<td>GABA A$\alpha$</td>
<td>AAGCCTCAGAGTCC TTCCAGTATG</td>
<td>2,090</td>
<td>AAATAAGTGJAAGA AAAAAATTAGAT</td>
</tr>
</tbody>
</table>

www.aacrjournals.org

Clinical Cancer Research 2005;11(10) May 15, 2005 3726
and B899P with mammaglobin resulted in specificities higher than that of the combination of mammaglobin with B305D-C, the overall accuracy was decreased due to their lower sensitivities.

The bootstrapped estimate of specificity for mammaglobin and B305D-C decreased by only 2.4% from the apparent value of specificity calculated on the data collected (Table 6). The bootstrapped estimate of sensitivity, however, decreased by 5.8% from the apparent value of sensitivity (Table 6). This indicates that we can expect to observe a similar specificity in future trials, but the sensitivity may be more variable. A probable contributor to the weakness of the sensitivity is the small proportion of breast cancer patients in an already small sample set.

**Discussion**

Early detection of breast cancer has the clear potential to improve the management of patients. Because a substantial proportion of breast cancer patients have cancer cells present in their blood at time of diagnosis (8, 24), breast cancer cells seem to be released into the circulation at an early stage of the disease (8). The combination of cell enrichment techniques and molecular technologies provides an avenue to isolate and characterize circulating tumor cells to potentially provide effective and practical diagnostic and prognostic tests (6, 7, 41).

Recent findings showed the prognostic significance of circulating tumor cell levels in a prospective trial of 177...
metastatic breast cancer patients (42). In this study, patients with higher levels of circulating tumor cells (>5 circulating tumor cells per 7.5 mL whole blood compared with <5), before treatment initiation, had a shorter median progression-free survival (2.7 versus 7.0 months, \( P < 0.001 \)) and shorter overall survival (10.1 versus >18 months, \( P < 0.001 \)). The Food and Drug Administration–approved CellSearch System (Veridex, Warren, NJ), including the CellSearch Epithelial Cell Kit and the CellSpotter Analyzer (43), was used for isolation and enumeration of circulating tumor cells that express the epithelial-cell adhesion molecule. This trial established the clinical utility of enumerating circulating tumor cells to predict survival of metastatic breast cancer patients.

In our study, we wanted to establish the diagnostic potential of circulating tumor cell characterization in patients with suspicious mammograms. Similar to Cristofanilli and coworkers, we used an immunomagnetic epithelial cell enrichment method. However, the procedure we used enriches the samples for cells expressing BerEP4 instead of enrichment method. However, the procedure we used coworkers, we used an immunomagnetic epithelial cell with suspicious mammograms. Similar to Cristofanilli and potential of circulating tumor cell characterization in patients with metastatic breast cancer patients.

In our study, we wanted to establish the diagnostic potential of circulating tumor cell characterization in patients with suspicious mammograms. Similar to Cristofanilli and coworkers, we used an immunomagnetic epithelial cell enrichment method. However, the procedure we used enriches the samples for cells expressing BerEP4 instead of epithelial-cell adhesion molecule. In addition, we determined the expression level of highly breast-specific genes in the captured epithelial cells using qRT-PCR for mammaglobin, B305D-C, B726P, GABA \( \alpha_{x} \), and CK-19. We then correlated these findings with biopsy results. Our results established the clinical feasibility of using molecular characterization of circulating tumor cells with mammaglobin and B305D-C for early detection of breast cancer. Our two-gene marker test (at select gene expression levels) correctly identified 78% of the patients with benign or invasive disease.

Similar to results from previous studies using qRT-PCR (15, 21), we observed a large range in the distributions of gene expression in circulating epithelial cells. Fortunately, qRT-PCR allowed us to construct receiver operator characteristic curves for each marker to differentiate between the baseline levels of gene expression in benign breast disease patients and the increased levels of gene expression associated with invasive breast cancer.

### Table 2. Patient characteristics

<table>
<thead>
<tr>
<th>Benign breast disease (( n = 100 ))</th>
<th>DCIS (( n = 10 ))</th>
<th>Invasive breast cancer (( n = 44 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median age at biopsy (range)</strong></td>
<td>51 (21-84)</td>
<td>62 (41-75)</td>
</tr>
<tr>
<td><strong>Nodal status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nx</td>
<td>9 (90%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>N−</td>
<td>1 (10%)</td>
<td>29 (66%)</td>
</tr>
<tr>
<td>N+</td>
<td>0 (0%)</td>
<td>14 (32%)</td>
</tr>
<tr>
<td><strong>Menopausal status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal</td>
<td>34 (34%)</td>
<td>1 (10%)</td>
</tr>
<tr>
<td>Perimenopausal</td>
<td>11 (11%)</td>
<td>1 (10%)</td>
</tr>
<tr>
<td>Postmenopausal natural</td>
<td>27 (27%)</td>
<td>4 (40%)</td>
</tr>
<tr>
<td>Postmenopausal surgical</td>
<td>28 (28%)</td>
<td>4 (40%)</td>
</tr>
</tbody>
</table>

**NOTE:** Positive nodal status (+) was determined as having at least one positive axillary node detected at the time of primary surgery.

The transcript levels of mammaglobin, B305D-C, B726P, GABA \( \alpha_{x} \), and CK-19 in circulating cells obtained from DCIS patients levels were not significantly different from those observed from patients with benign breast disease. This would be expected from the noninvasive nature of DCIS. In contrast, mammaglobin mRNA levels seemed to be higher in patients with invasive breast cancer than in DCIS patients. We also found that mammaglobin, B305D-C, and CK-19 transcript levels were significantly higher in patients with invasive breast cancer than in those with benign breast disease. Our results showed that circulating epithelial cells detected with mammaglobin, CK-19, and B305D-C occurred at a higher frequency in patients diagnosed with invasive breast cancer than in those with benign breast disease. These results suggest that invasive but not noninvasive tumors shed tumor cells into the circulation and support the clinical experience that DCIS is not metastatic. However, the statistical tests involving DCIS patients (\( n = 10 \)) have little power and the resulting conclusions must be interpreted with caution.

In contrast to previous findings that showed no correlation between CK-19 and mammaglobin gene expression in circulating cells from primary breast cancer patients (18), we

### Table 3. Characteristics of 44 patients with positive biopsies

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Histology</strong></td>
<td></td>
</tr>
<tr>
<td>Invasive ductal carcinoma</td>
<td>38 (86%)</td>
</tr>
<tr>
<td>Invasive lobular carcinoma</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>Invasive tubular carcinoma</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Not specified</td>
<td>1 (2%)</td>
</tr>
<tr>
<td><strong>Nodal status</strong></td>
<td></td>
</tr>
<tr>
<td>Well</td>
<td>13 (30%)</td>
</tr>
<tr>
<td>Moderate</td>
<td>16 (36%)</td>
</tr>
<tr>
<td>Poor</td>
<td>14 (32%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (2%)</td>
</tr>
<tr>
<td><strong>Size (cm)</strong></td>
<td></td>
</tr>
<tr>
<td>≤1</td>
<td>11 (25%)</td>
</tr>
<tr>
<td>1-2</td>
<td>16 (36%)</td>
</tr>
<tr>
<td>2-5</td>
<td>14 (32%)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (2%)</td>
</tr>
<tr>
<td><strong>ER status</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>34 (77%)</td>
</tr>
<tr>
<td>Negative</td>
<td>7 (16%)</td>
</tr>
<tr>
<td>Not determined</td>
<td>3 (7%)</td>
</tr>
<tr>
<td><strong>PR status</strong></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>33 (75%)</td>
</tr>
<tr>
<td>Negative</td>
<td>8 (18%)</td>
</tr>
<tr>
<td>Not determined</td>
<td>3 (7%)</td>
</tr>
<tr>
<td><strong>Her2 status</strong></td>
<td></td>
</tr>
<tr>
<td>IHC 0/1+</td>
<td>13 (30%)</td>
</tr>
<tr>
<td>IHC 2+</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>IHC 3+</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>Fluorescence in situ hybridization not amplified</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Not determined</td>
<td>24 (55%)</td>
</tr>
</tbody>
</table>

**NOTE:** A positive ER or PR status (+) is determined as having any neoplastic cell immunohistochemically stained for ER or PR.
observed significant associations between CK-19 and mammaglobin transcript levels in both benign breast disease and invasive breast cancer patients with a much stronger association in breast cancer patients ($r_s = 0.27$ versus $0.65$).

Similar to our results, however, previous findings showed that mammaglobin was superior to CK-19 as a marker for minimal residual disease of breast cancer (18, 29, 30). CK-19 has been extensively used as a marker for breast cancer cells but with limited success. Low levels of keratin gene expression are normally present in lymphoid and other nonepithelial cells (25, 44). In addition, CK-19 previously detected circulating epithelial cells in healthy individuals and negative controls (19). Thus, we focused this study on mammaglobin, B305D-C, B726P, and GABA $\alpha_2$, which previously showed promising diagnostic potential in breast cancer (12, 45, 46).

Numerous investigators have used mammaglobin to detect circulating cells in metastatic and nonmetastatic breast cancer patients using traditional RT-PCR techniques (11, 12, 18, 45–52). These earlier studies reported that mammaglobin was detected in circulating cells in over 50% of metastatic breast cancer patients (12, 48, 49, 51) and in 24% and 9% treated breast cancer patients with and without metastases, respectively (52). Presumably, the circulating tumor burden would be higher and more likely to be detected in metastatic breast cancer patients, especially before initiation of treatment.

In a study of patients more similar to our patient population, mammaglobin was detected in 24% of 78 pretreated, primary breast cancer patients (47). Mammaglobin also was detected in 8% of 111 invasive breast cancer patients (18). Similarly, we observed that at 100% specificity, 9% of the invasive breast cancer patients were correctly identified by mammaglobin. Further, at 75% specificity, mammaglobin correctly identified 64% of invasive breast cancer patients. We observed an increase in sensitivity compared with previous findings, probably because we used a cell enrichment procedure, a higher initial blood volume, and qRT-PCR.

Although little or no mammaglobin gene expression in control cell lines and normal, healthy individuals was observed in these previous studies, others have shown mammaglobin mRNA in 43% and 59% of nonmalignant and fibroadenoma breast tissue, respectively (53). In our study, we observed relatively high mammaglobin transcript levels (>46.3 units) in 14% of patients with benign biopsies. Because mammaglobin has been observed in nonmalignant breast tissue (11, 50, 51, 53) and our “negative” samples were actually patients who had benign breast disease, the significant background level observed in some patients may

---

**Table 4. Individual sensitivity and specificity of mammaglobin, B305D-C, B726P, GABA $\alpha_2$, and CK-19**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cut point (units)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Percentage correctly identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammaglobin</td>
<td>32.2</td>
<td>63.6 (48.9, 76.2)</td>
<td>75.0 (65.7, 82.5)</td>
<td>71.5</td>
</tr>
<tr>
<td>B305D-C</td>
<td>3.7</td>
<td>59.1 (44.4, 72.3)</td>
<td>60.0 (50.2, 69.3)</td>
<td>59.7</td>
</tr>
<tr>
<td>B726P</td>
<td>0.5</td>
<td>59.2 (44.4, 72.3)</td>
<td>58.0 (48.2, 67.2)</td>
<td>58.3</td>
</tr>
<tr>
<td>GABA $\alpha_2$</td>
<td>0.2</td>
<td>100.0 (92.0, 100.0)</td>
<td>0.0 (0.0, 3.7)</td>
<td>30.6</td>
</tr>
<tr>
<td>CK-19</td>
<td>42.1</td>
<td>59.1 (44.4, 72.6)</td>
<td>60.0 (50.2, 69.3)</td>
<td>59.7</td>
</tr>
</tbody>
</table>

NOTE: Values in parentheses represent the corresponding 95% confidence intervals. Cut points were determined by the intersection of receiver operator characteristic curves ($n = 144$). Units are in gene copies/pg $\beta$-actin.
be due to increased self-breast palpation after receiving the news of a suspicious mammogram. This increased self-palpation may stimulate the release of epithelial cells into the circulation. Alternatively, because certain histologic types of benign breast disease (e.g., fibroadenomas, atypical hyperplasia) may result in an increased risk of breast cancer in certain patients (54), one can speculate that the extremely high \textit{mammaglobin} levels observed in the few patients with benign breast disease may indicate an enhanced risk of later developing invasive breast cancer. Other possible reasons for high background \textit{mammaglobin} levels include illegitimate expression of the marker as often seen with keratins (25, 44, 55) or genomic DNA contamination from poor separation of the lysed cell suspension from the magnetic epithelial beads. In addition, preliminary results from our laboratory indicated that circulating epithelial cells were detected by \textit{mammaglobin} in healthy females. The \textit{mammaglobin} mRNA levels varied with respect to lactation, menstrual cycle, and menopausal status (this is under further investigation). In support of these observations, Watson et al. (50) mentioned unpublished data suggesting that \textit{mammaglobin} expression was associated with mammary gland proliferation and terminal differentiation. Also, high levels of \textit{mammaglobin} in breast cancer tissue were found at lower serum estradiol levels after menopause (56).

Although \textit{mammaglobin} is a promising tumor marker, it is not universally expressed in all breast cancers (11, 29, 45, 46, 50, 51, 53, 57). Our results showed that ~20% of invasive breast cancer patients did not have detectable levels of \textit{mammaglobin}. Therefore, we evaluated the utility of adding B305D-C, B726P, GABA $A_x$, or CK-19 to the analysis of \textit{mammaglobin} to discriminate between patients with benign and invasive breast cancer breast biopsies. Our results showed that combining \textit{mammaglobin} with B305D-C improved both sensitivity and specificity. In addition, \textit{mammaglobin} or B305D-C was detected in 86% of the N+ patients. These results support the findings that \textit{mammaglobin} tissue expression may be a marker of axillary lymph node breast metastases (57). More recent findings suggested that elevated \textit{mammaglobin} expression in breast tumors was associated with clinical and biological features defining a less aggressive tumor phenotype (56, 58).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cut point (units)</th>
<th>Gene</th>
<th>Cut point (units)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Percentage correctly identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammaglobin</td>
<td>46.3</td>
<td>B305D-C</td>
<td>11.49</td>
<td>70.5 (55.8, 81.8)</td>
<td>81.0 (72.2, 87.5)</td>
<td>77.8</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>46.3</td>
<td>B726P</td>
<td>29.5</td>
<td>56.8 (42.2, 70.3)</td>
<td>86.0 (78.7, 90.7)</td>
<td>76.4</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>46.3</td>
<td>B899P</td>
<td>18.3</td>
<td>59.1 (44.4, 72.3)</td>
<td>85.0 (75.6, 89.9)</td>
<td>76.4</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>46.3</td>
<td>CK-19</td>
<td>79.3</td>
<td>68.2 (53.4, 80.0)</td>
<td>75.0 (65.7, 82.5)</td>
<td>72.9</td>
</tr>
</tbody>
</table>

*NOTE: Values in parentheses represent the corresponding 95% confidence intervals (n = 144). Units are in gene copies/pg $\beta$-actin.*
Breast tumors are highly heterogeneous and, thus, a panel of several markers may be more reliable than a single target for sensitive detection of malignant breast cancer cells (7, 12, 15, 29, 30, 45, 46). To this end, we screened other cancer-related genes, including urokinase plasminogen activator (UPA), UPA-inhibitor-type 1, osteopontin, and human anterior gradient-2R. However, the addition of these genes did not add to the diagnostic potential of the assay. Other candidate genes include B511S, carcinoembryonic antigen, prolactin inducible protein, muc1, PS2, p1B, and EGP2, which showed prognostic potential in breast cancer (15, 21, 29, 30). Increased sensitivity also may result from using less stringent criteria for detection limits as we classified samples with a C, higher than 37 as nondetectable. Indeed, false-negative results occur when levels of molecular target are near the detection limit of the RT-PCR technique (59). Further, very recent findings showed that a combination of four tests, including cytology, fluorescence in situ hybridization for aneuploidy, and RT-PCR for mamoglobin A, and mamoglobin B, improved the diagnosis of malignant effusions and yielded a sensitivity and specificity of 81% and 70%, respectively (60). In contrast, our approach utilized only one methodology and yielded a sensitivity and specificity of 70.5% and 81.0%, respectively.

In addition to determining the levels of circulating tumor cells (42), gene expression analysis offers an opportunity for developing biomarker-based tests that are less expensive and more accurate than existing screening tests. Our results suggest that sensitive molecular techniques can be used alone or in combination with mammography and physical examination for early and accurate detection of invasive breast cancer. Additional markers and simple modifications to our procedure (e.g., negative cell enrichment, nested PCR) may improve sensitivity and specificity of the screening test. Further work with a larger population of patients is needed to validate this method as a beneficial screening test for breast cancer. Overall, the results from this feasibility study suggest that the molecular characterization of circulating tumor cells based on the detection of breast-specific genes, including mammaglobin and B305D-C, has excellent potential as an additional screening tool for breast cancer.

References

Evaluation of a Panel of Tumor Markers for Molecular Detection of Circulating Cancer Cells in Women with Suspected Breast Cancer

Monica M. Reinholz, Andrea Nibbe, Leslie M. Jonart, et al.


Updated version Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/11/10/3722

Cited articles This article cites 58 articles, 24 of which you can access for free at: http://clincancerres.aacrjournals.org/content/11/10/3722.full.html#ref-list-1

Citing articles This article has been cited by 10 HighWire-hosted articles. Access the articles at: /content/11/10/3722.full.html#related-urls

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.