Immunohistochemical and Clinical Characterization of the Basal-Like Subtype of Invasive Breast Carcinoma

Torsten O. Nielsen,[1] Forrest D. Hsu,[1] Kristin Jensen,[2] Maggie Cheang,[1] Gamze Karaca,[4,5] Zhiyuan Hu,[4,5] Tina Hernandez-Boussard,[3] Chad Livasy,[4,6] Dave Cowan,[4,7] Lynn Dressler,[4,7] Lars A. Akslen,[8] Joseph Ragaz,[9] Allen M. Gown,[10] C. Blake Gilks,[1] Matt van de Rijn,[2] and Charles M. Perou[4,5,6]

[1]Genetic Pathology Evaluation Centre, University of British Columbia, Vancouver Hospital & British Columbia Cancer Agency, Vancouver, British Columbia, Canada; Departments of [2]Pathology and [3]Genetics, Stanford University Medical Center, Stanford, California; [4]Lineberger Comprehensive Cancer Center and Departments of [5]Genetics, [6]Pathology and Laboratory Medicine, and [7]Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; [8]Department of Pathology, The Gade Institute, Haukeland University Hospital, Bergen, Norway; [9]McGill University Health Center, Royal Victoria Hospital, Montreal, Quebec, Canada; and [10]PhenoPath Laboratories, Seattle, Washington

ABSTRACT

Purpose: **Expression profiling studies classified breast carcinomas into estrogen receptor (ER)+/luminal, normal breast-like, HER2 overexpressing, and basal-like groups, with the latter two associated with poor outcomes. Currently, there exist clinical assays that identify ER+/luminal and HER2-overexpressing tumors, and we sought to develop a clinical assay for breast basal-like tumors.**

Experimental Design: **To identify an immunohistochemical profile for breast basal-like tumors, we collected a series of known basal-like tumors and tested them for protein patterns that are characteristic of this subtype. Next, we examined the significance of these protein patterns using tissue microarrays and evaluated the prognostic significance of these findings.**

Results: **Using a panel of 21 basal-like tumors, which was determined using gene expression profiles, we saw that this subtype was typically immunohistochemically negative for estrogen receptor and HER2 but positive for basal cytokeratins, HER1, and/or c-KIT. Using breast carcinoma tissue microarrays representing 930 patients with 17.4-year mean follow-up, basal cytokeratin expression was associated with low disease-specific survival. HER1 expression was observed in 54% of cases positive for basal cytokeratins (*versus* 11% of negative cases) and was associated with poor survival independent of nodal status and size. c-KIT expression was more common in basal-like tumors than in other breast cancers but did not influence prognosis.**

Conclusions: **A panel of four antibodies (ER, HER1, HER2, and cytokeratin 5/6) can accurately identify basal-like tumors using standard available clinical tools and shows high specificity. These studies show that many basal-like tumors express HER1, which suggests candidate drugs for evaluation in these patients.**

Received 2/4/04; revised 4/30/04; accepted 5/11/04.

Grant support: M. Cheang and C. Gilks were supported by an educational grant from Aventis, Inc. M. van de Rijn was supported by funds from National Cancer Institute (NCI) Grant CA85129. C. Perou was supported by funds from the NCI Breast Specialized Programs of Research Excellence Grant P50-CA58223-09A1 (University of North Carolina at Chapel Hill) and NCI Grant CA-101227-01.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: T. Nielsen is a Michael Smith Foundation for Health Research Scholar.

Requests for reprints: Charles M. Perou, Lineberger Comprehensive Cancer Center, CB# 7295, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599. Phone: (919) 843-5740; Fax: (919) 843-5718, E-mail: cperou@med.unc.edu.

INTRODUCTION

Recent DNA microarray profiling studies on breast tumors have identified distinct subtypes of breast carcinomas that are associated with different clinical outcomes (1, 2). Using an intrinsic set of 534 genes, Sørlie *et al.* (2) analyzed the expression profiles of 115 independent breast tumor samples and categorized breast tumors into five groups: luminal A [estrogen receptor (ER)+]; luminal B (ER+); HER2 overexpressing; normal breast-like; and basal-like. On the basis of >300 tumors with expression profiles and associated clinical follow-up data spanning three independent data sets, breast cancers of the basal-like subtype comprised 19% of the tumors and had poor prognoses as assessed by relapse-free survival (2–4). Therapies targeting the ER or HER2 oncogene would not be expected to be effective on basal-like breast cancers because this subtype typically expresses neither of these proteins. Although diagnostic antibodies that work in formalin-fixed, paraffin-embedded archival tissue do not exist for most genes present in the basal-like breast cancer gene expression profile, commercial antibodies to cytokeratin 5/6 and cytokeratin 17 are available. The prevalence and poor prognosis of basal-like breast cancers has been validated immunohistochemically on a 564 case tissue microarray (TMA) with 66-month mean outcome data using overall survival as an end point (5); 16% of tumors in this cohort stained positive for cytokeratin 5/6 and/or cytokeratin 17.

Recently, another independent TMA study of basal cytokeratin expression and related immunohistochemical markers was published (6). In this study, breast cancers that were cytokeratin 5/6 positive were found to be associated with expression of the epidermal growth factor receptor (HER1), with the proliferation marker Ki-67, with accumulation of p53 and

with increased cytogenetic abnormalities. In another recent study, the basal-like subtype, as defined by cytokeratin 5/6 expression by immunohistochemistry (IHC), was also found to be common among breast cancer patients with hereditary BRCA1 mutations (7).

Basal-like breast cancers represent a poorly characterized subtype of tumor with no validated clinical assay to identify them; therefore, in this report, we first improved our immunohistochemical definition of basal-like breast cancers by comparison to our gene expression data. Next, based on a larger cohort of breast cancer patients with longer follow-up, we show that the basal-like breast cancer tumors show poor disease-specific survival times and show that HER1 expression is a marker that helps to distinguish basal-like breast cancers. These data identify a simple set of IHC markers that can be routinely used in the clinical setting to accurately identify basal-like breast cancers.

MATERIALS AND METHODS

DNA Microarrays. The microarray data from Sørlie *et al.* (2) were used to identify genes/proteins whose pattern of expression could assist in identifying basal-like breast cancers using IHC. These data encompassed expression values for 8700 genes obtained from 115 grossly dissected tumors and also contained associated clinical information. Three additional basal-like breast cancers were identified at University of North Carolina at Chapel Hill using microarray analysis; total RNA from these tumors was isolated using a Qiagen RNeasy Midi kit according to the manufacturer's protocol (Qiagen, Valencia, CA). Next, we performed microarray analysis by first labeling the samples using an Agilent Low Input Linear Amplification kit (Agilent Technologies, Palo Alto, CA) and 2 μg of starting total RNA. These experiments used Agilent Human A1 oligo microarrays. The tumor samples were labeled with Cy5 and were compared with a common reference sample that was composed of the Stratagene Universal Human Reference sample that was augmented with a 1/10 amount of RNA from the MCF7 cell line and a 1/10 amount of RNA from the immortal human mammary epithelial cell line ME16C (8). The microarrays were scanned on an Axon 4000B Scanner with image analysis accomplished using GenePix Pro 4.0. The raw data (.gpr files) tables were uploaded into the University of North Carolina Microarray Database, which is a mirror of the Stanford Microarray Database (9). A global, linear normalization was performed to adjust the Cy3 and Cy5 channels (9). The three new microarray raw data tables are available at the University of North Carolina Microarray Database[11] and have been deposited into the Gene Expression Omnibus under the accession numbers GSM21709, GSM21710, and GSM21711.

TMAs. Cases for this study were drawn from 2475 patients participating in British Columbia Cancer Agency trials conducted between the late 1970s and 1990, involving stage I–III breast cancer (10–13). Cases for the microarray were selected based on the availability of paraffin blocks and represent a subset of patients in each trial. Paraffin blocks of formalin-fixed tissue were available for 930 patients. All patients have information on the date of diagnosis, age at diagnosis, date and type of relapse, and (in all but five cases) date and cause of death; most patients also have complete information on tumor size and histology as well as nodal status (14). The study was approved by the Clinical Research Ethics Board of the University of British Columbia. Patients have been followed regularly, with the last update in 2001 (mean and median follow-up from original diagnosis until analysis for patients still alive at time of analysis is 17.4 years; range, 9.8–28.1 years).

The TMA was constructed by extracting 600-μm diameter cores of histologically confirmed invasive breast carcinoma from the original paraffin blocks using a Beecher Instruments tissue core extractor and re-embedding these cores into a gridded paraffin block. Three such recipient blocks were constructed, containing 333, 334, and 336 tissue cores each arranged in four 12×7 sectors (15, 16). Control tissue cores from benign breast and kidney were included in each of these paraffin blocks. Single cores were taken from the original biopsy tissue from each of 930 patients. Forty-two additional cores were available from patients who had additional surgeries. After construction, 4-μm tissue sections were cut and adhered to Fisher SuperFrost Plus glass slides.

IHC and Scoring. Each set of three glass slides comprising the TMA was stained with commercially available antibodies: ER (Ventana, Tucson, AZ) antibody was used at 1:50 dilution, with a 4-min microwave antigen retrieval in citrate buffer. Cytokeratin 5/6 (Boehringer Mannheim, Indianapolis, IN) and cytokeratin 17 (DAKO, Carpinteria, CA) antibodies were both used at 1:10 dilution with antigen retrieval by proteinase K and microwave as above. c-KIT (also known as CD117 and as SCFR) antibody A4502 (DAKO) was used at 1:200 dilution with antigen retrieval by 40 min of steam treatment in EDTA buffer. HER1 and HER2 were stained using the PharmDX and Herceptest kits, respectively, according to manufacturer's instructions (DAKO). Detection was by EnVision+ (DAKO) with diaminobenzidine chromogen as per routine protocol. Five-μm whole tissue sections from 21 surgical cases selected because of their basal-like gene expression profile were treated similarly. ER status for the 115 samples used in the expression profiling studies of Sørlie *et al.* (2) was previously published and based on a ligand binding assay. For the 18 basal-like breast cancers examined here from that study, we also performed IHC for ER.

Staining results were assessed by at least two pathologists, using a three-point scoring system, where 0 = invasive tumor cells present in the tissue core and no staining seen, 1 = invasive tumor cells present with weak staining intensity and/or $<20\%$ of tumor cells stained, and 2 = invasive tumor cells present with strong staining in $>20\%$ of tumor cells. A positive ER stain was recorded only if immunostaining was seen within the nuclei of invasive carcinoma cells, whereas positive HER2 required strong membranous staining. Cytokeratin 5/6, cytokeratin 17, and HER1 were scored positive if any (weak or strong) cytoplasmic and/or membranous invasive carcinoma cell staining was observed. c-KIT immunostaining with polyclonal antibody A4502 had a higher degree of nonspecific background than the other antibodies used and was interpreted such that a score of 1 required 25–75% of cells positive and 2 required $>75\%$ cells

[11] Internet address: https://genome.unc.edu/.

Table 1 Frequency of immunostaining among arrayed breast carcinoma cases

Antigen	Interpretable cores	Positive staining (%)	Negative staining (%)
ER	755	302 (40)	453 (60)
HER2	744	150 (20)	594 (80)
CK5/6	788	110 (14)	678 (86)
CK17	765	35 (5)	730 (95)
HER1	614	82 (13)	532 (87)
c-KIT	761	105 (14)	656 (86)

strongly positive for cytoplasmic and/or membranous staining. Tissue cores that failed to adhere to the glass slide, did not contain invasive carcinoma, had been cut through, or were otherwise uninterpretable were excluded, with the numbers of informative cases for each marker shown in Table 1. Primary immunostained slide image data for ER, HER2, cytokeratin 5/6, cytokeratin 17, c-KIT, and HER1 were digitally acquired using a Bacus Laboratories, Inc., Slide Scanner system (Lombard, IL), and these images (~5000) are available online.[12]

Statistical Analysis. Results were tabulated into a Microsoft Excel worksheet format using the TMA-Deconvoluter program (17) and exported into the SPSS 11.0 statistical suite. For data analysis, immunostained cores scored as either 1 or 2 were considered positive, except in the case of HER2, where a score of 2 was required to be considered positive (equivalent to strong 3+ staining when using the Herceptest). Kaplan-Meier survival analyses were carried out for both overall and breast cancer disease-specific survival, using the Breslow test for differences between groups (18), where the proportional hazards assumption was not violated by crossing hazard function curves; log-rank statistics were also calculated. Prognostic variables were analyzed using univariate analysis, paired analysis, and jointly by stepwise as well as by overall regressions. Multivariate analyses was performed where the proportional hazards assumption was not violated, using a Cox regression model. Results were considered statistically significant when P from a two-tailed test was <0.05. Hierarchical clustering of TMA immunostaining results was performed as described in Nielsen *et al.* (19), with filters set to 70% interpretable staining data/core.

RESULTS

Comparison of cDNA Microarray and IHC Patterns in Basal-Like Tumors. Our gene expression profiling data divided breast carcinomas into five subtypes: luminal A and B tumors (both clinically ER+); HER2 positive; normal breast-like; and basal-like (1, 2, 20). A goal of this study was to develop a clinically applicable assay to identify breast basal-like cancers. Therefore, we first performed a review of our gene expression data for basal-like breast cancers from Sørlie *et al.* (2) to determine what genes/proteins might be exploitable for the clinical categorization of these tumors. Our analysis of the 115 tumors from Sørlie *et al.* (2) revealed that the basal-like breast cancers as a group (19 of 115) showed high expression levels for cytokeratin 5, HER1, and c-KIT and had low to absent gene expression of ER and HER2 (Fig. 1*A*). To determine whether this finding could be used to identify basal-like breast cancers, we asked if patients were first selected to be ER− and HER2−, which we define as a tumor that was not HER2 3+ by IHC and was <10 fmol/mg using a ligand-binding assay for ER, then how many of these patients had basal-like breast cancers. Of the 115 tumors assayed in Sørlie *et al.* (2) by microarray, 72 had clinical data for ER and HER2; from these 72 tumors, 18 were clinically ER− and HER2−. Starting with these 18 ER−/HER2− tumors, we determined that 15 were basal-like breast cancers by microarray analysis, suggesting that the majority (15 of 18) of clinically determined ER−/HER2− tumors were basal-like breast cancers.

A second way to characterize the basal-like subtype is to identify these tumors using microarray analysis and then determine which IHC patterns correlate with this expression profile. As shown in Fig. 1*A*, in the 115 tumors taken from Sørlie *et al.* (2), the basal-like breast cancers (red dendrogram branch) showed high expression of cytokeratin 5 and also tended to show high expression of HER1 and/or c-KIT. To corroborate these mRNA findings, we obtained paraffin sections from 18 of the basal-like tumors from Sørlie *et al.* (2), and from 3 more basal-like breast cancers from our current microarray studies, which were identified using the breast intrinsic gene list and hierarchical clustering analysis, giving us 21 tumors that were basal-like breast cancers as defined by DNA microarray analysis. We next performed IHC on these 21 tumors and other breast tumors for cytokeratin 5/6, HER1, c-KIT, ER, and HER2 (Fig. 1*B*). Examples of immunostaining for cytokeratin 5/6 and HER1 on some of the basal-like tumors and a normal breast sample are presented in Fig. 1*C*. Cases positive for ER showed the characteristic nuclear staining pattern (data not shown), whereas cytokeratin 5/6-positive cases showed a cytosolic and often focal pattern in tumors; in normal breast, cytokeratin 5/6 staining has been shown to display a wide variety of staining patterns from identifying the myoepithelial cell layer to identifying a layer of cells that line the ducts and that may represent committed stem cells (21). HER2-positive cases showed a membranous pattern (data not shown), whereas HER1 (Fig. 1*C*) and c-KIT-positive tumors showed a mixed cell membranous and cytoplasmic pattern; nuclear/nuclear membranous staining was not seen. In normal breast, c-KIT staining was present in most epithelial cells, whereas HER1 staining was largely restricted to the stromal cell side of the basal/myoepithelial cell layer (Fig. 1*C*).

Using these IHC reagents, we determined that 13 of 21 basal-like breast cancers were cytokeratin 5/6 positive, 12 of 21 were HER1 positive, and 6 of 21 were c-KIT positive (Fig. 1*B*) when we scored both weak positive and strong positive cases as positive. These data suggest that an IHC surrogate for gene array experiments to identify basal-like breast cancers is to select for cases that are ER-negative, HER2-negative-low, and cytokeratin 5/6+ and/or HER1+. All 16 tumors with this immunohistochemical profile had the basal-like gene expression profile (Fig. 1). In total, 21 tumors were basal-like by DNA microarray, meaning the IHC surrogate definition is 76% sensitive and 100% specific. An alternative to this assay that some are using

[12] Internet address: http://microarray-pubs.stanford.edu/tma_portal/her1/index.shtml.

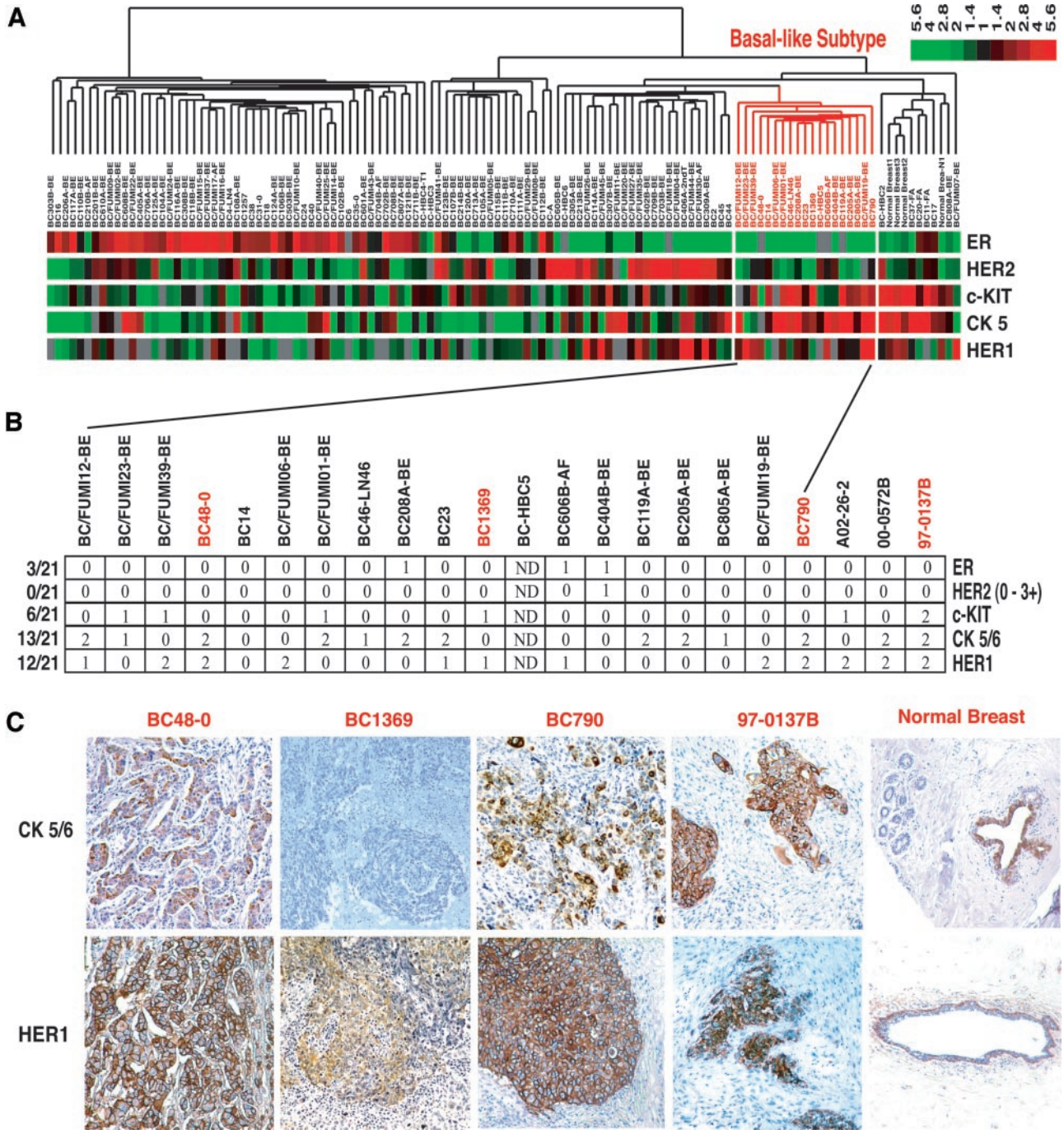

Fig. 1 Gene expression patterns in basal-like tumors and their correlation with immunohistochemistry. *A*, the 115 patient/tumor sample dendrogram was taken from the hierarchical clustering analysis presented in Sorlie *et al.* (2); the tumors were grouped using the breast intrinsic gene list based on 45 paired samples. The basal-like breast cancers are identified in *red*. The gene expression data for ER, HER2, c-KIT, cytokeratin (CK)5, and HER1 are shown with *red squares* representing the highest average expression, *black* representing average gene expression, and *green* representing the lowest below average. *B*, 21 basal-like breast cancers identified by gene expression profiling were tested and scored by IHC for CK5/6, HER1, c-KIT, ER, and HER2 (0 = negative, 1 = weak and/or focal staining, 2 = strong diffuse staining), except for HER2, which was scored using a standard (0–3+) scale. ND = not determined. *C*, representative immunostaining results for four basal-like tumors and a normal breast sample for CK5/6 and HER1.

to identify basal-like breast cancers is to select for ER-negative and HER2-negative-low tumors; however, for technical reasons, it can be dangerous to base an assay upon the absence of all staining.

Association of Basal Cytokeratins with Poor Clinical Outcome. We previously demonstrated that expression of basal cytokeratin markers by IHC in breast carcinomas (cytokeratin 5/6+ and/or cytokeratin 17+) predicted poor outcomes (5).

To corroborate these data, we used a different patient cohort TMA containing 930 cases with a median follow-up time of 17.4 years. The overall frequency of staining observed for each of the immunohistochemical markers over the interpretable cases is presented in Table 1. Kaplan-Meier survival analysis of cases with an interpretable score for cytokeratin 5/6 and/or cytokeratin 17 ($n = 829$) shows that positivity for either of these basal markers correlates with shorter disease-specific survival than for negative cases (Fig. 2*A*; median survival 8.8 *versus* 13.2 years, $P = 0.015$). From this set of patients, nodal status was available for 682, of whom 472 were positive and 210 negative for lymph node metastasis. In the lymph node-positive group, the presence of either basal cytokeratin was associated with a significantly poorer outcome (Fig. 2*B*; $P = 0.008$ by Breslow test); in the lymph node-negative group, a trend was seen but did not reach statistical significance (Fig. 2*C*). These data confirm our earlier finding (Ref. 5 and see "Discussion"); however, because the staining pattern of cytokeratin 17 is often difficult to score, we sought to improve upon this classification.

Association of the Basal-Like Subtype with HER1 and c-KIT Expression. Part of our new basal-like breast cancers immunoprofile was expression of HER1. On the current 930 case TMA series, HER1 expression was present in 44.1% (41 of 93) of the cancers that were positive for a basal cytokeratin, which was similar to the percentage seen on the DNA microarray cohort described above (12 of 21 = 57%). HER1 expression was significantly less common among the basal cytokeratin-negative cases (41 of 521 = 7.9%, $P < 0.001$ by Fisher's exact test) and was present in only 2.0% (5 of 254) of ER-positive cases. HER1 was demonstrable in a much larger fraction (82 of 614 = 13.4%) of breast cancers than cytokeratin 17 (35 of 765 = 4.6%), suggesting that inclusion of HER1 into the IHC definition of basal-like tumors might be more valuable than cytokeratin 17. By comparison, cytokeratin 5/6 immunostaining was positive in 110 of 788 = 14.0% of the breast cancers tested. In a univariate analysis, HER1 expression was associated with poor survival (median, 5.7 *versus* 13.4 years; $P < 0.01$ by log-rank and Breslow) as shown in Fig. 2*D*. In multivariate analysis, expression of HER1 was a significant independent negative prognostic factor ($P = 0.017$, relative risk 1.54) when fitting the available clinical variables of nodal status (relative risk 2.10) and tumor size (relative risk 1.12) in a Cox proportional hazards model. Histological grade could not be used in this multivariate analysis because accurate grading information was unavailable for the majority of cases.

c-KIT expression was present in 14% (105/761) of all breast cancers when strong and weak staining intensities were both considered positive. 31% (32/102) of the cancers positive for basal cytokeratins were also positive for c-KIT, whereas only 11% (67/605) of basal cytokeratin negative cases were positive ($P < 0.001$ by Fisher's exact test); in addition, we determined by IHC that 6 of 21 basal-like breast cancers from the microarray studies were also c-KIT positive (Fig. 1). When only strong c-KIT staining was considered, only 2.2% (17 of 761) of the tumors were positive. This is precisely the frequency obtained by Simon *et al.* (22), who determined that 43 of 1654 patients in their breast tumor TMA study expressed c-KIT; Simon *et al.* (22) also showed that none of these tumors contained mutated c-KIT. By mRNA expression, c-KIT is one of the best basal-specific markers; however, by IHC, many of the

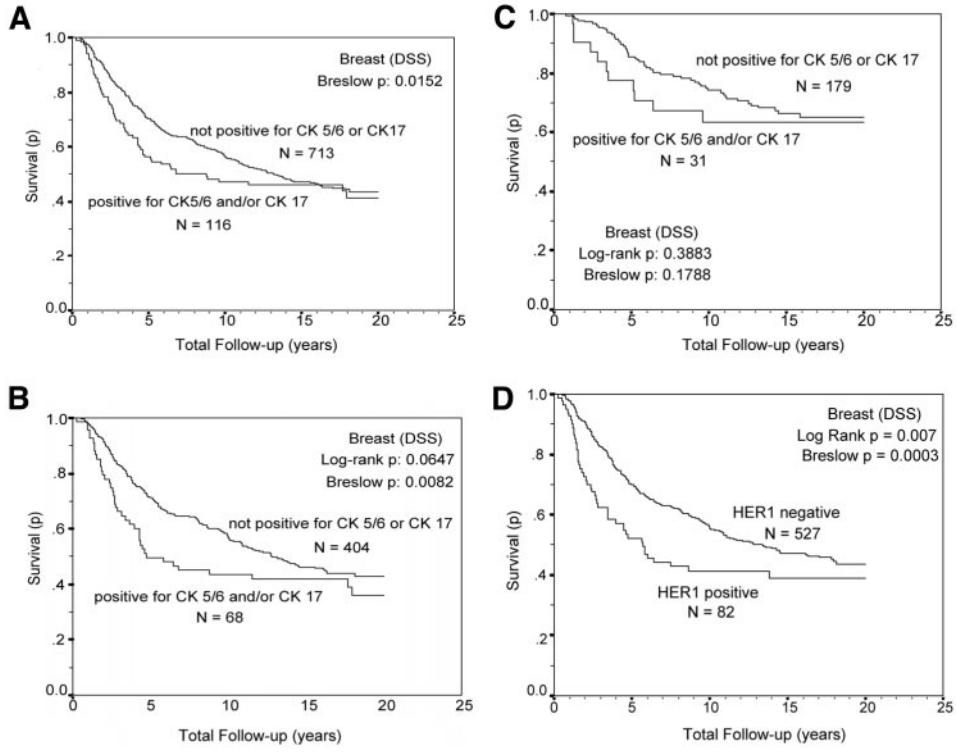

Fig. 2 Kaplan-Meier survival curves based on basal cytokeratin and HER1 staining. *A*, Kaplan-Meier disease-specific survival (DSS) curve for 829 tumors assessed for cytokeratin (CK)5/6 and CK17. *B*, Kaplan-Meier DSS curve for 472 lymph node-positive breast cancers assessed for cytokeratin 5/6 and cytokeratin 17. *C*, Kaplan-Meier DSS curve for 210 lymph node-negative breast cancers assessed for CK5/6 and CK17. *D*, Kaplan-Meier DSS curve in 609 breast cancers assessed for HER1 immunostaining.

tumors that showed c-KIT mRNA were not positive for c-KIT protein expression. In contrast to HER1, c-KIT was not associated with significant differences in patient outcome on our TMA cohort regardless of whether positivity was called only for strong immunostaining, or for strong and weak combined (data not shown).

Refinement of an IHC Definition for Basal-Like Tumors. Given the association of HER1 with the basal subtype, its prognostic significance in univariate analysis and its greater sensitivity when compared with cytokeratin 17, we sought to incorporate this marker into an immunohistochemical definition of basal-like breast cancers. A total of 663 TMA core samples had sufficient interpretable staining results to allow sample characterization into one of four groups (Table 2). These categories were defined in a clinically practical way: (*a*) if a tumor is HER2 positive (*i.e.*, 3+ by IHC), it falls immediately into group H; (*b*) if a tumor is HER2 negative-low and ER positive, then it is group E; (*c*) if a tumor is both HER2 and ER negative but positive for at least one basal-enriched marker (cytokeratin 5/6 and/or HER1), then that tumor falls into group B; and (*d*) if a tumor is negative for all four markers, it falls into group N (undetermined). Using this definition, the basal-like breast cancers (group B) comprised 15% (102 of 663) of the tumors studied. Fig. 3 shows a hierarchical cluster diagram of the TMA results for the six protein markers and shows the correlation in expression among the basal-like breast cancer-enriched markers cytokeratin 5/6, cytokeratin 17, HER1, and c-KIT. A version of this cluster diagram exists on our web site;[12] at this web site, digital images are available for all tissue cores by clicking on the StainFinder field next to each individual tumor.

Basal-like breast cancers, defined immunohistochemically as ER negative, HER2 negative, and cytokeratin 5/6 and/or HER1 positive, not only correlated best with the basal-like breast cancers gene expression profile but also gave the most significant survival differences on TMA (Fig. 4) with a much worse prognosis than for group E (ER+). Outcomes were similar to those of the HER2+/group H; of note, these data were collected before the use of trastuzumab therapy for HER2+ tumors. These results agree with the clinical outcomes of basal-like breast cancers in the Norway/Stanford and Dutch breast tumor expression profiling studies (1–3), which further validates the clinical significance of the basal-like breast cancers in a fashion that can be applied to standard formalin-fixed, paraffin-embedded tissue. Tumors negative for all markers (group N) also fared poorly and probably include tumors that represent a number of different subtypes; likely included in this uncategorized group are a few basal-like breast cancers not detected by cytokeratin 5/6 or HER1 because of their imperfect sensitivity and some-

Fig. 3 Hierarchical cluster analysis of breast carcinoma TMA immunostaining results. The six markers evaluated in this study [ER, HER1, HER2, c-KIT, cytokeratin (CK)5/6, and CK17] were scored and used in a clustering analysis where each row represents a different tumor, and each column represents a different IHC stain. The analysis shows that the basal-like markers stained similar cases, as indicated by the very short dendrogram branches linking these markers. A more detailed version of this figure, linked to the primary TMA image files, is available online.[12]

Table 2 Frequencies of the immunohistochemically defined subtypes of breast cancers in 663 carcinoma cases informative for the tested markers using tissue microarrays

Group	HER2	ER	Cytokeratin 5/6 and/or HER1	Frequency (%)
H (HER2)	Positive	Any	Any	150 (23)
E (luminal)	Negative	Positive	Any	263 (40)
B (basal-like)	Negative	Negative	One or both positive	102 (15)
N (negative)	Negative	Negative	Negative	148 (22)

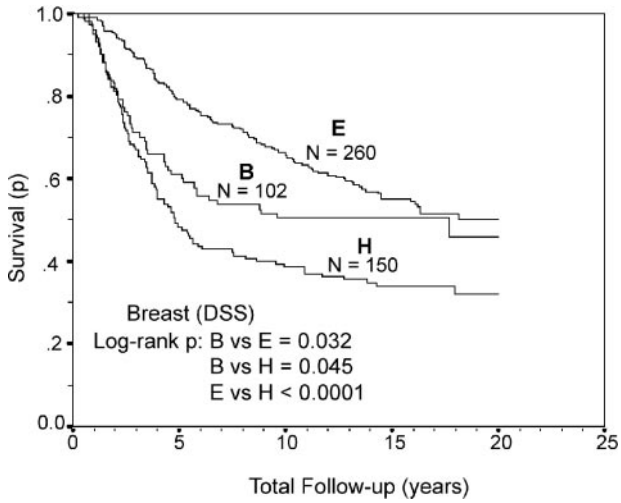

Fig. 4 Kaplan Meier disease-specific survival (DSS) analysis of breast carcinoma subtypes defined by ER, HER2, CK5/6, and HER1 immunohistochemistry. Group B (the basal-like subtype) is defined as negative for ERs, negative for HER2, and positive for HER1 and/or CK5/6. Group E (the luminal subtype) are positive for ERs and negative for HER2. Group H (clinically HER2 3+) are positive for HER2.

times focal expression pattern, luminal-derived tumors that very weakly express ER, normal breast-like tumors, and tumors put into this category due technical failures. Because of the heterogeneity within the all-negative group, at least one positive marker is needed to assure high specificity, and it is for this reason that our assay includes cytokeratin 5/6 and/or HER1 positivity. Finally, the basal-like breast cancers that were HER1 positive but cytokeratin 5/6 negative could differ in outcomes from their counterparts who were HER1 negative and cytokeratin 5/6 positive. However, the HER1-positive only basal-like breast cancers and the cytokeratin 5/6-positive only basal-like breast cancers showed equally poor outcomes, both with statistical significance in a univariate analysis (data not shown).

DISCUSSION

Building upon our gene expression studies (1, 2, 20), van de Rijn *et al.* (5) found that expression of the basal-like breast cancers markers cytokeratin 5/6 and cytokeratin 17 predicted poor outcome in breast tumor patients. Using a similar approach here, our current TMA study was able to again validate the clinical importance of the basal-like breast cancers subtype. The current and larger series allowed confirmation of the frequency (15%) of this subtype on an independent cohort and found a clear association with short disease-specific survival, as was also seen in other studies (2, 23). We did identify one difference in outcomes between our two TMA studies, which was that in the van de Rijn *et al.* (5) study, the cytokeratin 5/6 and/or cytokeratin 17-positive tumors were significant predictors of outcome in the node-negative subset, whereas in this study, the cytokeratin 5/6- and/or cytokeratin 17-positive tumors were a significant predictor in the node-positive group; we believe this finding may be due to the different patient cohorts and are examining this point on additional cohorts.

Bocker *et al.* (24) and Korsching *et al.* (6) have hypothesized that a basal-like stem cell, characterized by its preferential expression of cytokeratin 5/6 and low expression of luminal cytokeratins (cytokeratin 8/18), might represent a breast cancer subclass. Wetzels *et al.* (25) found a strong correlation between basal cytokeratin expression and cell proliferation, which was corroborated by our profiling studies (1, 20, 26). However, the staining pattern of the basal keratins (cytokeratin 5/6 and especially cytokeratin 17) is challenging to detect by immunohistochemical methods because of focal and often weak reactivity. On the basis of a review of our gene expression data followed by immunohistochemical validation, we found that we can better define the basal-like breast cancers by identifying those tumors that are negative for both ER and HER2 and that are positive for cytokeratin 5/6 and/or HER1. The HER1 marker is much easier to score than cytokeratin 5/6 and is much more frequent than cytokeratin 17. Also, the use of a single basal marker (cytokeratin 5/6), whereas successful in identifying a subset of patients with poor outcomes, misses approximately half of basal-like tumors. In addition, reliance on the lack of staining for ER and HER2 alone to identify basal-like breast cancers risks misassignment based on technical failures and/or biological heterogeneity.

HER1 expression is not a basal-like breast cancer-specific marker like cytokeratin 5/6; however, it was expressed in enough basal-like breast cancers that when combined with other markers, it greatly assists in their immunohistochemical identification. More importantly, HER1 is also a target for several recently developed drugs, including therapeutic antibodies (cetuximab) and small molecule tyrosine kinase inhibitors (gefitinib, erlotinib; Refs. 27, 28). The evaluation of HER1 inhibitors as a monotherapy in unselected breast cancer patients has begun; however, the association of HER1 expression with basal-like breast cancers could define a subset of breast cancer patients who might benefit from treatment with one of these agents, either alone, or in combination with standard chemotherapy. Our results also show a relationship between c-KIT expression and the basal-like breast cancer subtype, with the majority of c-KIT-positive breast tumors belonging to the basal-like breast cancer subtype. Here, we show that a simple panel of four antibodies can robustly identify basal-like breast cancers using standard IHC, which could serve as the basis to identify basal-like breast cancer patients in the clinical setting and for the retrospective evaluation of the efficacy of known chemotherapeutic agents on this tumor subtype.

ACKNOWLEDGMENTS

We thank William Gerald for contributing paraffin sections from a number of basal-like tumors and we acknowledge the technical assistance and support of the Tissue Procurement and Analysis Facility of University of North Carolina at Chapel Hill.

REFERENCES

1. Sørlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869–74.

2. Sørlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci USA 2003;100:8418–23.

3. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature (Lond.) 2002;415:530–6.

4. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA 2003;100:10393–8.

5. van de Rijn M. Perou CM, Tibshirani R, et al. Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. Am J Pathol 2002;161:1991–6.

6. Korsching E, Packeisen J, Agelopoulos K, et al. Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. Lab Investig 2002;82:1525–33.

7. Foulkes WD, Stefansson IM, Chappuis PO, et al. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J Natl Cancer Inst (Bethesda) 2003;95:1482–5.

8. Troester MA, Hoadley KA, Sorlie T, et al. Cell-type-specific responses to chemotherapeutics in breast cancer. Cancer Res 2004;64: 4218–26.

9. Sherlock G, Hernandez-Boussard T, Kasarskis A, et al. The Stanford microarray database. Nucleic Acids Res 2001;29:152–5.

10. Ragaz J, Olivotto IA, O'Reilly S, et al. The significance of mastectomy in patients with locally advanced breast cancer treated with preoperative (neoadjuvant) therapy: is there a need for randomization? Proc Am Soc Clin Oncol 1991;10:41a.

11. Ragaz J, Baird R, Rebbeck P, Goldie J, Coldman A, Spinelli J. Neoadjuvant (preoperative) chemotherapy for breast cancer. Cancer (Phila.) 1985;56:719–24.

12. Ragaz J, Jackson SM, Le N, et al. Adjuvant radiotherapy and chemotherapy in node-positive premenopausal women with breast cancer. N Engl J Med 1997;337:956–62.

13. Ragaz J, Coldman AJ, Le N, Olivotto I, Fang R, Hislop TG. Incidence of second cancers in patients with primary breast cancer: analysis of BC Cancer Registry cohorts with differing policies of adjuvant therapy between 1970–1994. Proc Am Soc Clin Oncol 1998;17: 421a.

14. Nielsen TO, Andrews HN, Cheang M, et al. Expression of the insulin-like growth factor-1 receptor and urokinase plasminogen activator in breast cancer is associated with poor survival: potential for intervention with 17-allylamino geldanamycin. Cancer Res 2004;64: 286–91.

15. Schraml P, Kononen J, Bubendorf L, et al. Tissue microarrays for gene amplification surveys in many different tumor types. Clin Cancer Res 1999;5:1966–75.

16. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–7.

17. Liu CL, Prapong W, Natkunam Y, et al. Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. Am J Pathol 2002;161:1557–65.

18. Breslow NE. A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. Biometrika 1970;57: 579–94.

19. Nielsen TO, Hsu FD, O'Connell JX, et al. Tissue microarray validation of epidermal growth factor receptor and SALL2 in synovial sarcoma with comparison to tumors of similar histology. Am J Pathol 2003;163:1449–56.

20. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature (Lond.) 2000;406:747–52.

21. Boecker W, Buerger H. Evidence of progenitor cells of glandular and myoepithelial cell lineages in the human adult female breast epithelium: a new progenitor (adult stem) cell concept. Cell Prolif 2003;36 (Suppl 1):73–84.

22. Simon R, Panussis S, Maurer R, et al. KIT (CD117)-positive breast cancers are infrequent and lack KIT gene mutations. Clin Cancer Res 2004;10:178–83.

23. Malzahn K, Mitze M, Thoenes M, Moll R. Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas. Virchows Arch 1998;433:119–29.

24. Bocker W, Bier B, Freytag G, et al. An immunohistochemical study of the breast using antibodies to basal and luminal keratins, alpha-smooth muscle actin, vimentin, collagen IV and laminin. Part I: normal breast and benign proliferative lesions. Virchows Arch A Pathol Anat Histopathol 1992;421:315–22.

25. Wetzels RH, Kuijpers HJ, Lane EB, et al. Basal cell-specific and hyperproliferation-related keratins in human breast cancer. Am J Pathol 1991;138:751–63.

26. Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. Nat Genet 2002;32 (Suppl):533–40.

27. Baselga J. Targeting the epidermal growth factor receptor with tyrosine kinase inhibitors: small molecules, big hopes. J Clin Oncol 2002;20:2217–9.

28. Baselga J, Hammond LA. HER-targeted tyrosine-kinase inhibitors. Oncology 2002;63 (Suppl 1):6–16.

AACR American Association for Cancer Research

Clinical Cancer Research

Immunohistochemical and Clinical Characterization of the Basal-Like Subtype of Invasive Breast Carcinoma

Torsten O. Nielsen, Forrest D. Hsu, Kristin Jensen, et al.

Clin Cancer Res 2004;10:5367-5374.

Updated version Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/10/16/5367

Cited articles This article cites 24 articles, 8 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/10/16/5367.full#ref-list-1

Citing articles This article has been cited by 100 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/10/16/5367.full#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, use this link
http://clincancerres.aacrjournals.org/content/10/16/5367.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC)
Rightslink site.

Downloaded from clincancerres.aacrjournals.org on February 27, 2018. © 2004 American Association for Cancer Research.