Award Number: DAMD17-01-1-0463

TITLE: Outcome Based Screening for Prognostic Phospho-RTK (Receptor Tyrosine Kinase) Antibodies Using Tissue Microarrays

PRINCIPAL INVESTIGATOR: David L. Rimm, M.D., Ph.D.

CONTRACTING ORGANIZATION: Yale University
New Haven, Connecticut 06520-8047

REPORT DATE: August 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
**Report Title:**
Outcome Based Screening for Prognostic Phospho-RTK (Receptor Tyrosine Kinase) Antibodies Using Tissue Microarrays

**Authors:**
David L. Rimm, M.D., Ph.D.

**Performing Organization:**
Yale University
New Haven, Connecticut  06520-8047

E-Mail: david.rimm@yale.edu

**Sponsoring Agency:**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland  21702-5012

**Supplementary Notes:**
Original contains color plates. All DTIC reproductions will be in black and white.

**Abstract:**
Receptor Tyrosine Kinases (RTKs) have been identified as potential targets for both breast cancer prognosis and therapy. We proposed use of tissue microarrays to evaluate the prognostic value of RTKs with emphasis on the phosphorylation status of these receptors. Analysis of a series of phospho-receptor antibodies on large cohorts on tissue microarrays should reveal which RTKs are most likely to be of prognostic and therapeutic value. As of this progress report, we have completed construction of the tissue microarrays and completed collection of the clinical data. We have also completed and submitted a pilot study of RTKs using conventional analysis of this array. In the proposal we show automated analysis of the arrays has potential to reveal relationships that are undetectable by conventional methods. We have now completed our efforts in construction of the device capable of high through-put automated array analysis. We have begun using this device with some phospho-specific antibodies to EGFR and Met.
# Table of Contents

Cover................................................................................................................. 1  
SF 298............................................................................................................... 2  
Introduction.................................................................................................... 4  
Body.................................................................................................................. 4  
Key Research Accomplishments................................................................. 7  
Reportable Outcomes..................................................................................... 7  
Conclusions..................................................................................................... 8  
References....................................................................................................... 8  
Appendices...................................................................................................... 9  

Introduction:

Biologically specific therapies represent a great new hope for combating cancer. Perhaps the best example of this is Herceptin, a specific drug for a subset of patients with HER2 positive breast cancer. Unfortunately, the pathway from initial discovery to clinical usage is long and slow, taking over 10 years for HER2 (a member of the receptor tyrosine kinases (RTK) family. In this study we proposed a potential method for rapid evaluation of other RTKs as bio-specific therapies based on prediction of outcome. We proposed that phospho-specific antibodies to RTK would provide highly specific prognostic markers for outcome and predictive markers for response to anti-receptor type therapies. Our objective was to produce specific phospho-RTK antibodies and then to evaluate them using breast cancer tissue microarrays. Breast cancer tissue microarrays are a method of placing 0.6 mm diameter samples of breast cancers from hundreds of patients on a single slide. Using this method we proposed high throughput screening for potential prognostic antibodies at the earliest stages of antibody development.

Body:
The main difficulty in finding these potential new RTK targets is that there are many potential candidates and it is difficult and expensive to evaluate a large number of antibodies on large breast cancer cohorts using conventional methods. The novelty or “idea” of this proposal is the use of a newly described, high throughput mechanism for evaluation of antibodies. Thus instead of producing candidate antibodies and screening using the conventional approach (which can take many years), we propose reversal of this process, screening for candidates using the outcome testing, then only proceeding with development of antibodies that have already been validated on large populations of breast cancer patients. We originally proposed the following specific aims:

2. Evaluation of mouse test bleeds using 250 case, 3-fold redundant, breast cancer cohort tissue microarrays.
3. Selection of promising sera and production of monoclonal antibodies followed by confirmatory testing on tissue microarrays.

Our original Statement of Work was as follows:

Year 1:

1. Select and produce phospho- and corresponding non-phospho-peptides representing critical sequences of RTKs
2. Select 250 breast cancer cases and begin construction of tissue microarray
3. Inoculate mice with first set of phospho-peptides and collect test bleeds.
4. Begin first array screening

Year 2

1. Collect test bleeds and screen arrays.
2. Do final inoculations and test bleeds
3. Begin full scale TMA verification of first promising candidates
4. Do fusions and begin production of first Mabs
5. Production of small scale cloning verification tissue microarrays

Year 3
1. Complete production of Mabs
2. Complete verification of Mabs using first small scale TMAs
3. Production and large cohort testing (750+ case) of new antibodies

This progress report describes work completed in years 1 and 2.

In evaluation of the optimal RTKs for antibody preparation we consulted many sources. We found that many biotech antibody companies had already embarked on production of phosphospecific and other RTK antibody production. Rather than try to duplicate their efforts, we decided to begin by purchasing some of the 100’s of RTK antibodies that are now commercially available.

Prior to testing any antibodies, arrays needed to be constructed and cohort data collected. That process was completed during the first year. We now have constructed 3 master array blocks for analysis. They include YTMA10, including 350 node positive cases, YTMA 12, including 350 node negative cases, and YTMA 23, including a total of 250 cases representing a subset of the previous cohorts (125 node negative and 125 node positive) that have ample tissue available for analysis of multiple antibodies.

After completion of construction of the arrays, we began a two-fold approach to analysis. As a proof of concept, we selected 5 RTKs (HER2, MET, EGFR, FGFR, and IGFR) as test cases. Dr. Tolgay Ocal, a pathologist recruited to this effort, analyzed expression of these using YTMA 12. He found that only Met was predictive of outcome in the node-negative population but that we were able to group MET and FGFR in a single group that was unrelated to HER2 and EGFR which comprised a second, unrelated group. Furthermore, he found that the localization of the epitope on Met was critical for prognostication of outcome. Specifically, the C-terminal antibody (3D4) was significantly associated with poor outcome, but a second antibody recognizing the extra-cellular domain of the Met beta chain was not at all associated with outcome. This was confirmed in a second effort using the same array looking at Met, matriptase, HAI 1 and HGF led by Julie Kang. Both of these efforts have now been published and both papers are included in the appendix.

Although we were pleased with these results, the original grant discussed analysis of arrays using a more quantitative method than the ordinal subjective scale used in this study. Toward that aim, other investigators in the lab have finalized a method for automated quantitative analysis of tissue microarrays. This work will be extremely valuable in optimizing the high throughput analysis of arrays in the future. Although the automated analysis development work was not funded by this grant, it impacts the grant dramatically. The paper that describes the new automated analysis system (AQUA) was published last November and is also included in the appendix.
The automated analysis system was designed on the Applied Precision Deltavision platform, but that platform is not well suited to the task and the system at our institution is very heavily used allowing only limited time for our applications. Approval of a rebudgeting request allowed the purchase of a new dedicated 2nd generation automated tissue microarray analysis device. This device (called Professor Marvel, in honor of its location behind the curtain), has now been operational for about 9 months.

The next step was the evaluation of multiple phospho-specific antibodies. Although we originally proposed making our own phospho-specific antibodies, since that proposal, many have become commercially available. Thus we have purchased a number of antibodies to the tyrosines proposed in the original grant. To date we have focused on the phospho-tyrosines of Met, EGFR, and HER2. The antibodies have been purchased from Cell Signaling Technologies (CST), Upstate Biochemical and Biosource. We have now collected data on all of the phospho-tyrosines on EGFR using the 250 case TMA and are well on our way to collecting data on phospho-Met. These antibodies have shown some unexpected subcellular localizations (specifically many are nuclear). This has led to questions regarding antibody cross reactivity or non-specificity.

We have begun with validation studies of three Phospho-Met antibodies from BioSource. Although these are not monoclonals, they are made from three phospho-peptides (Y1003, Y1235, and Y1349), and then affinity purified. Others and we have shown these antibodies work on western blots by showing increased phosphorylation after stimulation with HGF, the Met ligand. However, the patterns on test arrays were variable from case to case and the localization was not similar to that seen with some Met antibodies. Toward validation of these antibodies, first we used incubation with phosphatases to determine if the antigens are destroyed. We found that Calf Intestinal Alkaline Phosphatase was sufficient to abolish essentially all activity of the phospho-Met antibodies. Next we tested the specificity of each of the three phospho-met antibodies by using purified phospho-peptides to assess competitive inhibition of the interactions on tissue arrays. Figure 1 shows that the Y1349 antibody is somewhat diminished (row 3) or barely at all diminished (row 1) by the phospho-Y1349 peptide, but it is similarly competed by the phospho-Y1003 and Y1235 peptides. We conclude that even though there is specificity on a western blot, on fixed tissue there is significant cross-reactivity that will prevent future use of these antibodies in this application.

Although were are still concerned about the heterogeneity of phosphorylation, we plan to proceed to test further phospho-specific monoclonal antibodies, beginning with the Cell Signaling Technologies phospho-EGFR antibodies. Depending on the results of these studies, we anticipate efforts for further validation and progression to the VEGFR and FGFR receptors in the final year of the grant. The data collection has largely been done by Julie Kang and, more recently, by Matt Neapolitano. Drs. Rimm and Ocal are doing the analysis of the data.
Figure 1: Competition assays of phospho-specific antibodies by phospho-peptides show non-specific peptides are as effective as specific peptides for competition in immunofluorescence assays. Rows 1 and 3 are the antibody to Y1349 of the cytoplasmic domain of Met, shown at 20X and 40X respectively. Row 2 is the Y1003 antibody. Row 1 shows the phospho-peptide used for each column of images: column 1, none; column 2, Y1003; column 3, Y1234; column 4 Y1349.

Key Research Accomplishments:

- Completion of technology for high throughput automated analysis of tissue microarrays
- Completion of array construction of both the 250 case array and the 700 case array and cohort
- Completion of RTK conventional pathologist-based analysis of array with node negative cohort
- Testing of phospho-specific antibodies on arrays

Reportable Outcomes:

Analysis of RTKs shows MET is predictive of outcome in a node-negative population and it identifies a series of patients unique from that identified by over-expression of HER2 or EGFR but related to those that over-express FGFR.
Conclusions:

Tissue microarrays are a valuable tool for analysis of protein and phospho-protein expression. Completion of our automated analysis system will allow high throughput analysis of these arrays. We anticipate evaluation of the validated phospho-specific RTK antibodies within the next year.

The significance of these developments is that we are now well positioned to find a series of valuable prognostic markers. These markers will are likely to be valuable for prognostication, but they are also likely to be valuable in that they have the potential to identify pathways that will be good targets for bio-specific therapeutics.

References:
See original proposal.
Appendices:


Automated subcellular localization and quantification of protein expression in tissue microarrays

ROBERT L. CAMP, GINA G. CHUNG & DAVID L. RIMM
Department of Pathology, Yale University School of Medicine, New Haven, Connecticut, USA
Correspondence should be addressed to D.L.R.; email: david.rimm@yale.edu

Published online 21 October 2002; doi:10.1038/nm791

The recent development of tissue microarrays—composed of hundreds of tissue sections from different tumors arrayed on a single glass slide—facilitates rapid evaluation of large-scale outcome studies. Realization of this potential depends on the ability to rapidly and precisely quantify the protein expression within each tissue spot. We have developed a set of algorithms that allow the rapid, automated, continuous and quantitative analysis of tissue microarrays, including the separation of tumor from stromal elements and the sub-cellular localization of signals. Validation studies using estrogen receptor in breast carcinoma show that automated analysis matches or exceeds the results of conventional pathologist-based scoring. Automated analysis and sub-cellular localization of beta-catenin in colon cancer identifies two novel, prognostically significant tumor subsets, not detected by traditional pathologist-based scoring. Development of automated analysis technology empowers tissue microarrays for use in discovery-type experiments (more typical of cDNA microarrays), with the added advantage of inclusion of long-term demographic and patient outcome information.

Despite the promise of automated analysis of histological sections, it has failed to replace traditional, pathologist-based evaluation, even in the simplest of conditions such as the analysis of immunohistochemical stains. Whereas the automated analysis of isolated cells in fluids or smears (for example, fluorescent cell sorting and laser scan cytometry) is now routine, the analysis of tissue sections is hampered by the fact that tumor tissue is a complex mixture of overlapping malignant tumor cells, benign host-derived cells and extracellular material. Several methods (including confocal and convolution/de-convolution microscopy) can determine the subcellular localization of target antigens, but only through computationally intensive techniques, requiring the acquisition of multiple high-power, serial images. Methods designed for tissue microarrays perform only limited subcellular localization using morphometry and usually require significant manual interface (for example, drawing polygons around tumor cells). In general, pathologist-based analysis remains the current standard for the immunohistochemical studies.

Tissue microarrays provide a high-throughput method of analyzing the prognostic benefit of a myriad of potential targets on large cohorts of patient samples, but are limited by the pathologist's ability to reproducibly score on a continuous scale, discriminate between subtle low-level staining differences, and accurately score expression within subcellular compartments. We have developed a set of algorithms that we call AQUA (Automated Quantitative Analysis) that allow the rapid, automated analysis of large-scale cohorts on tissue microarrays. The first algorithm, called PLACE (pixel-based locale assignment for compartmentalization of expression) utilizes fluorescent tags to separate tumors from stroma and to define subcellular compartments. The distribution of a target antigen is then quantitatively assessed according to its co-localization with these tags. As subcellular compartments (for example, membrane, cytoplasm, nuclei, and so forth) of different tissues and tumors vary widely in size and shape, traditional methods of defining compartments based on morphometric criteria (that is, feature extraction) perform poorly on a large-scale basis. Rather than counting target-containing features, PLACE delineates target expression as the sum of its intensity divided by the total size of the assessed compartment.

As the thickness of tissue sections makes it difficult to discriminate between overlapping subcellular compartments, we have also developed a novel, rapid exponential subtraction algorithm (RESA), which subtracts an out-of-focus image, collected slightly below the bottom of the tissue, from an in-focus image, based on pixel intensity, signal-to-noise ratio, and the expected compartment size. This algorithm dramatically improves the assignment of pixels to a particular subcellular compartment (Fig. 1). For a more complete discussion of the image manipulations performed in this protocol, see Supplementary Methods or http://www.yalepath.org/dept/research/YCCTMA/tissarray.htm.

Validation of AQUA algorithms

Our initial validation of this technology compared its accuracy, intra-observer variability, and predictive power to traditional pathologist-based analysis. We stained a tissue microarray derived from 340 node-positive breast-carcinoma patients for the presence of estrogen receptor (ER)—the oldest and most common prognostic marker for breast cancer. First we analyzed the ability of automated analysis to match results from a pathologist-based evaluation and found a high degree of correlation (R = 0.884, Fig 2a). Next, we compared the variability of a pathologist-based and automated analysis of two separate histospots derived from the same tumor (Fig. 2b and c). This comparison shows that automated analysis has slightly better reproducibility (R = 0.824 versus R = 0.732).

Although automated analysis compares favorably with pathologist-based interpretation of microarrays, the true criterion standard is outcome prediction. Estrogen receptor expression is known to significantly improve outcome, because it is associated with less aggressive tumors that are more responsive to anti-estrogens (for example, Tamoxifen). We compared the survival of patients with tumors with high (top 25%) versus low (bottom 25%) ER expression as assessed by both automated and pathologist-based scoring (Fig. 2d). Results show that both methods provide similar prognostic information (RR = 2.44 versus 2.66, automated versus pathologist); although the auto-
NEW TECHNOLOGY

Fig. 1 RESA allows the accurate assignment of subcellular compartments and localization of a target antigen. A, A pseudo-three-color image of a colon carcinoma shows a significant degree of overlap between subcellular compartments: Blue, nuclei (DAPI); green, tumor mask (cytokeratin); red, tumor cell membranes (alpha-catenin). B, The signal intensity of a target antigen, beta-catenin (inset), is redistributed according to the relative signal intensity of the compartments identified in A: Blue, nuclear-localized; red, membrane-localized; green, cytoplasmic. Note that the beta-catenin expression in this tumor is predominantly membrane-associated, yet there is significant incorrectly assigned signal in the nucleus: magenta and blue pixels. C, The compartment-specific signals in A are reassembled using the RESA algorithm, reducing the amount of overlapping signal by exponentially subtracting pixel intensity from an out-of-focus image. D, The signal intensity from an exponentially subtracted image of the target antigen, beta-catenin (inset) is then redistributed according to the compartments defined in C. This result more accurately assign the target antigen to the membrane compartment (red pixels) with little expression in the nuclear compartment (blue pixels).

Univariate analysis shows slightly higher significance (P = 0.0003 versus P = 0.0020). Univariate analysis of the automated analysis shows a relative risk of 2.438 (P = 0.0005, 95% CI 1.480–4.016). When analyzed in a multivariate analysis against histological and nuclear grades, age and stage, automated ER analysis retains independent prognostic significance (RR = 2.566, 95% CI 1.428–4.611, P = 0.0016). The pathologist-based analysis shows similar results, validating the cohort (see Supplementary Methods).

To determine the reproducibility of our automated analysis of ER, we used the 'split-sample technique,' by dividing the cohort into halves and using one half as a 'training' set and the other as a 'test' set. The training set was used to determine standard cutoffs for the top and bottom 25% of cases. These cutoffs were then used to divide the test set into top, middle and bottom groups. We analyzed 300 randomly selected training and test sets; on average 97% of the test cases were correctly classified.

One clear advantage to automated analysis is that it can perform a true continuous assessment of a target. In contrast, the human eye, even that of a trained pathologist, has a difficult time accurately distinguishing subtle differences in staining intensity using a continuous scale. Consequently, scoring systems for pathologists tend to be nominal (for example, 0, 1+, 2+, 3+). Algorithms such as the 'H-score' are meant to translate such nominal observations into semi-quantitative results. However, the inability to detect subtle differences in staining intensity, particularly at the low and high ends of the scale, as well as the tendency to round scores limits the effectiveness of the H-score. The discontinuity of pathologist-based scoring, despite the use of an H-score algorithm, is exemplified in the

Fig. 2 Automated and pathologist-based scoring of ER shows a high degree of correlation and equal power in predicting outcome. A, Two replicate tissue microarrays of 345 node-positive breast carcinomas were stained and scored by a pathologist using a traditional H-score algorithm. Note that despite good correlation (R = 0.732), there is significant discontinuity in the scoring pattern with numerous scores clustered around the high and low-end of the scale. B, Automated analysis was performed using the RESA and PLACE algorithms and compared to the pathologist's score with a high degree of correlation (R = 0.884). C, A comparison of two automated reads of two different microarrays derived from the same cohort also shows excellent correlation (R = 0.824). D, To determine the potential of each system to predict outcome, survival curves were plotted for the top and bottom 25% ER expressers using each scoring system: Red, automated; blue, pathologist-based; dotted, top 25%; solid, bottom 25%. Both systems provided a statistically significant survival difference (P = 0.0003 and P = 0.0020, automated and pathologist-based, respectively).
ER staining results in Fig. 2. Note the preponderance of scores at 0, 100, 200 and 300. Furthermore, on average, over half of the cases were assigned to one extreme or the other (39% at 0 and 12% at 300). Thus 51% of the cases could not be effectively ranked. In contrast, the range of scores from the automated analysis is continuous from 0 to 1000. We hypothesize that the two key advantages of automated assessment, continuity of scoring and accurate subcellular localization, will allow tumor classification beyond that attainable by current methods.

Compartmental analysis of beta-catenin expression
To demonstrate this potential, we analyzed beta-catenin expression in colon cancer. Beta-catenin is an ideal candidate in that it exhibits complex subcellular localization and manifests oncogenic properties upon localization to the nucleus. Numerous studies have shown that beta-catenin plays a dual role in both cell-cell adhesion and cell proliferation, depending on its location. Membrane-associated beta-catenin stabilizes cadherin-mediated adhesion by facilitating the cytoskeletal attachment of adhesion complexes. In contrast, nuclear-associated beta-catenin activates several genes important in cell proliferation and invasion. In development, translocation of beta-catenin to the nucleus results from Wnt-mediated cell signaling. However, spurious activation of this pathway is often seen in tumors through mutation of beta-catenin or other proteins involved in its activation and/or degradation. Studies on the prognostic value of beta-catenin have been mixed.

The complex biology and uncertain prognostic value of beta-catenin made it a suitable candidate for assessing the value of quantitative subcellular localization. We studied a cohort of 310 colon cancers, using both pathologist-based and automated systems for scoring overall, nuclear and membrane-associated levels of beta-catenin expression. Manual analysis used a traditional 4-point nominal scale (0 through 3), whereas automated analysis used a continuous 1,000-point scale. In a previous study using a similar cohort, we were unable to find a prognostic value in assessing nuclear beta-catenin levels. These data were confirmed in our present study when comparing tu-
NEW TECHNOLOGY

mos expressing the highest levels of nuclear β-catenin (3+, representing 19% of the cases) versus the rest (Fig 3a) (P = 0.2354). We hypothesized that with the benefit of automated, continuous assessment, these 3+ cases could be subdivided into cases expressing very high versus high levels. We began by analyzing the top 25% of tumors expressing nuclear β-catenin, as assessed by automated analysis. This group shows a trend toward poorer survival (Fig 3b) (P = 0.0760). When we subset the tumors to assess the top 10% expressors, there is a statistically significant survival difference (Fig. 3c) (P = 0.0332, relative risk = 1.740). Further fractionation of the data reveals that the top 6.76% (15th percentile) exhibit even poorer survival with a higher statistical significance (Fig. 3d) (P = 0.0038, relative risk = 2.415). This analysis demonstrates the power of continuous automated assessment to define subsets of tumors not seen using standard pathologist-based assessment. Unlike β-catenin, subdividing ER into smaller and smaller subsets does not significantly alter its prognostic ability, suggesting that ER may be a truly continuous marker where no subpopulations exist.

Tumor classification using the AQUA algorithms

We then attempted a tumor classification based on comparative subcellular localization. As the translational of β-catenin from the nucleus to the membrane is thought to correlate with transcriptional activation, we analyzed the ratio of nuclear to membrane-localized β-catenin. By its nature, this type of analysis is essentially impossible without continuous scoring. A crude measurement of overall β-catenin levels using either a pathologist-based or an automated system fails to demonstrate a significant difference in survival between the highest and lowest expressing tumors (Fig. 4a and b, P = 0.9425 and P = 0.4551, respectively). In contrast, when we ratio the level of nuclear/membrane β-catenin, we find that tumors with a high ratio have a worse outcome than tumors with a low ratio (RR = 1.718, P = 0.0284). Note that this method defines a relatively large subset (25%) of tumors with poor prognostic, the majority of which are not identified by analyzing individual subcellular compartments. Indeed, a comparison of the tumors with the highest nuclear/membrane ratio (top 25%) versus the highest overall nuclear levels of β-catenin (top 25%) shows that there is only 47% overlap between the two subsets. Multivariate analysis of nuclear/membrane β-catenin ratios shows independent prognostic significance when analyzed with depth of invasion, nodal status, tumor grade and patient age (RR = 1.865, 95% CI = 1.068–3.259, P = 0.0285). In contrast, multivariate analysis of total nuclear β-catenin levels fail to show independent prognostic significance because it is highly correlated with nodal metastases (see Supplementary Methods online).

The methods presented here are highly adaptable to a number of tumor types and target markers. In most cases, compartment-specific tags are identical regardless of tumor type (DAPI for nuclei, cadherin/catenin complexes for membranes). Because the methods do not use heuristics models that require recognition of compartments according to size, shape or texture, they are fully adaptable to tumors with overlapping or pleomorphic cells and/or nuclei. Furthermore, the algorithms can be easily expanded to cover novel compartments or even virtual compartments (for example, mitochondria, lysosomes, cortical actin ring) or tumor types (for example, mesothelioma), as long as tags can be identified for the prospective compartments/cell types. In addition to ER and β-catenin, we have used these techniques to successfully analyze dozens of markers, including growth factor receptors, intracellular signaling molecules, and proliferation markers, using both conventional and phospho-specific antibodies (see Supplementary Methods online). Analysis of these targets requires only a standard antibody titration.

We have found that most antigens benefit from subcellular localization. Localization can be simple, such as determining the amount of the proliferation marker Ki-67 in tumor nuclei, or more complex as in the case of β-catenin. Localization of intracellular signaling molecules (for example, STATs) may be vital in assessing their potential as prognostic markers. In addition, recent studies have shown that membrane-bound growth factor receptors (for example, epidermal growth factors, EGFR and ERB-B4, and fibroblast growth factor) can translocate to the nucleus and may act as transcriptional regulators.

Subcellular localization of such markers may be critical to their use as prognostic markers in cancer.

Depending upon the array size, and the complexity of the compartmentalization, analyses using our current device take from 1–3 hours for image acquisition, and 1–2 hours for analysis. In our laboratory, the average pathologist-based analysis rate is 50–100 spots per hour and usually is performed in several sessions. To increase precision, two or more pathologists read the same array independently and then together to resolve discrepancies. Aside from being more accurate and more robust, automated analysis can be performed continuously and results tabulated immediately. We estimate that a fully integrated tissue microarray reader could be 30 to 50 times faster than pathologist-based scoring.

Methods

Tissue microarray design and processing. Paraffin-embedded formalin-fixed specimens from 345 cases of node-positive breast carcinoma (1962–1977) and 310 cases of colon carcinoma (1971–1982) were obtained, as available, from the archives of the Yale University Department of Pathology. Microarray slides were prepared, processed and stained as described in the Supplementary Methods online. For manual analysis, slides were visualized with diaminobenzidine (DAB). For automated analysis, slides were visualized with Cy-5 tyramide.

Image and data analysis. Monochromatic images of tissue microarray histospots were obtained using fluorescently labeled compartment specific tags (anti-cyto-, keratin, DAPI, α-catenin) as well as target signals (ER and β-catenin). Regions of tumor were identified using a mask derived from a ubiquitously expressed epithelial-specific antigen (either cytokeratin or α-catenin). Images were analyzed using RESA and PLACE algorithms as detailed in the Supplementary Methods online. Results were expressed as the intensity of the target signal in each compartment divided by the compartment. For ER, only nuclear-localized signal was used; for β-catenin total signal, the ratios of nuclear-to-membrane signal and nuclear-to-total signal were analyzed. Overall survival analysis was assessed using Kaplan–Meier analysis and the Mantel–Cox log-rank score for assessing statistical significance. Relative risk was assessed using the univariate and multivariate Cox-proportional hazards model.
Our data show that quantitative, continuous-scale, compartmentalized automated analysis of tissue microarrays can provide a rapid assessment of prognostic-based subsets in a variety of tumor markers that cannot be attained using pathologist-based techniques. Automated analysis is better able to discern subtle differences in staining intensity, particularly at the upper and lower extremes, which can distinguish novel prognostic associations. Furthermore, analysis of the subcellular distribution of certain signals, using the PLACE and RESA algorithms may elucidate previously unrecognized associations with patient survival. The automated nature of this technology can allow high-throughput screening of tissue microarrays, facilitating their use in large-scale, high-throughput applications such as target discovery and prognostic marker validation. If, someday, diagnostic criteria are based on molecular expression patterns, the digital nature of this analysis could allow a device of this type to make specific molecular diagnoses.

Note: Supplementary Information available on the Nature Medicine website.

Acknowledgments

We thank T. D'Aquilla, M. Helo, L. Charette, D. Fischer, E. Rimm and P. Lizardi for their help in this effort, and J. Costa, V. Marchesi, A. Reynolds, R. Levenson and E. Fearon for review of the manuscript. This work was supported by grants from the Patrick and Catherine Weldon Donaghue Foundation for Medical Research and grants from the NIH including: KO-8 ES11571, NIEHS (to R.L.C.), RO-1 GM57604 NCI (to D.L.R.) and US Army DAMD grant 01-00456.

Competing interests statement

The authors declare competing financial interests: see the website (http://nature.com/naturemedicine)

Tissue Microarray Analysis of Hepatocyte Growth Factor/Met Pathway Components Reveals a Role for Met, Matriptase, and Hepatocyte Growth Factor Activator Inhibitor 1 in the Progression of Node-negative Breast Cancer

Jung Y. Kang, Marisa Dolled-Filhart, Idris Tolgay Ocal, Baljit Singh, Chen-Yong Lin, Robert B. Dickson, David L. Rimm, and Robert L. Camp

Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06520 (J. Y. K., M. D.-F., I. T. O., D. L. R., R. L. C.), and Lombardi Cancer Center, Departments of Oncology and Pathology, Georgetown University School of Medicine, Washington, DC 20007 (B. S., C.-Y. L., R. B. D.)

ABSTRACT

Numerous studies have demonstrated that overexpression of Met, the hepatocyte growth factor (HGF) receptor, plays an important role in tumorigenesis. Met activation can either occur through ligand-independent or dependent mechanisms, both of which are mediated by a series of proteases and modulators. We studied the protein expression of several components of the HGF/Met pathway on a cohort of 330 node-negative breast carcinomas using a tissue microarray annotated with 30-year, disease-specific patient follow-up data. We examined HGF, matriptase (an activator of HGF expressed on mammary epithelial cell surfaces), HAI-1 (the natural inhibitor of matriptase), and the Met receptor itself. Our studies demonstrate tight correlation between the expression of HGF, matriptase, and Met in breast carcinoma. High-level expression of Met, matriptase, and HAI-1 were associated with poor patient outcome. Met and HAI-1 showed independent prognostic value when compared with traditional breast markers in a multivariate analysis. Intriguingly, antibodies against the intracellular but not the extracellular domain of Met were prognostic, suggesting that overexpression of the cytoplasmic tail of Met, perhaps through cleavage or truncating mutation, may play an important role in breast cancer progression.

INTRODUCTION

Many studies have demonstrated the importance of the HGF³ pathway in carcinogenesis (1). HGF is produced both by tumor cells as well as by surrounding stromal elements, and can act in either a paracrine or autocrine fashion (2–5). HGF is secreted as an inactive propeptide, which must be cleaved to become biologically active. One enzyme responsible for this cleavage is matriptase, an epithelial-localized transmembrane serine protease (6–8). Matriptase is in turn, regulated by a naturally occurring inhibitor, HAI-1 (9, 10). When cleaved, HGF can bind to its receptor, Met, thereby stimulating multiple downstream pathways, leading to mitogenesis, motogenesis, and morphogenesis (11).

Several studies have analyzed individual components of the HGF pathway for their association with tumor aggression and/or patient survival. Early biochemical studies demonstrated that overall levels of HGF in breast cancers correlated with worse patient outcome (2, 3, 5). Whether HGF production by tumors and/or surrounding stroma is an important prognostic feature is unclear; although tumor cells themselves are a major producer of HGF (4, 5, 12). The use of matriptase and HAI-1 as prognostic markers in breast cancer has not been reported previously. However, recently, one study demonstrated that high matriptase and low HAI-1 levels were associated with advanced-stage ovarian tumors (13). Another report demonstrated that the glycosylation of matriptase stabilized and enhanced its proteolytic activity, and promoted tumor aggression (14).

The expression of Met has been more extensively studied. Met overexpression associates with poor prognosis in a variety of tumors (1). Whether such expression is ligand-(HGF) dependent or independent is unclear; however, the constitutive activation of Met, via several ligand-independent mechanisms, is established. These mechanisms include activating point mutations (15–19), chromosomal translocations (20, 21), and truncations of the cytoplasmic domain (22, 23). In addition, dysregulation of Met-associated phosphatases may also lead to Met activation (24).

We have now studied several elements of the HGF pathway including HGF, matriptase, HAI-1, and Met in a single cohort of node-negative breast cancer patients with 30-year follow-up, correlating the expression of each element and determining their prognostic value. This study was facilitated by the use of tissue microarrays: arrays of hundreds of patient histological samples on a single glass slide. Our study demonstrates a significant correlation between members of the HGF pathway and shows that several members have independent prognostic value in determining patient outcome.

MATERIALS AND METHODS

Cohort Design and Tissue Microarray Construction. Tissue microarrays were constructed, as described previously, and reviewed recently (25, 26). Three hundred and thirty cases of formalin-fixed, paraffin-embedded, node-negative breast carcinoma were obtained from the archives of the Department of Pathology, Yale University. Cases were taken sequentially, as available, from 1962 to 1980, with a median survival time of 15.6 years. Complete treatment history is not available from this cohort, but the vast majority of patients in this era were not treated with chemotherapy. Representative tumor regions were selected for coring by a pathologist (R. L. C.). Because prior studies have demonstrated that a single core adequately represents the staining pattern of an entire slide, all of the studies were performed using a single sample of each tumor (27, 28). Our previous study has also demonstrated the durability of antigens from archival specimens as old as 70 years (27). In the present study, all of the tumors demonstrated some degree of staining with one or more of the antibodies tested, demonstrating that no cases were antigenically “dead” because of fixation artifacts or tissue age.

Immunohistochemistry. Briefly, 5-μm tissue microarray slides were deparaffinized with xylene and ethanol. Antigen retrieval was performed using citrate buffer (pH 6.0) pressure-cooking (29). Primary antibodies were incubated overnight at 4°C, with the exception of antibodies to ERs, PRs, and Her2, which were incubated at room temperature for 1 h. Monoclonal anti-matriptase
and anti-HAI-1 antibodies were prepared as described previously (9, 30). Commercially acquired antibodies included: polyclonal (goat) anti-HGF antibody (R&D Systems, Minneapolis, MN); monoclonal antibody to the extracellular domain of Met (DO-24; Upstate Biotechnology, Lake Placid, NY); and monoclonal antibody to the intracellular domain of Met (3D4; Zymed, South San Francisco, CA). The specificities of all of the antibodies used were verified using immunoprecipitation and Western blotting. Antibodies to ER, PR, and HER-2/neu were obtained from DAKO (Carpinteria, CA) and used according to the manufacturer's specifications. Antibodies were either detected using a Vectastain ABC kit (Vector Laboratories, Burlingame, CA) for anti-HAI-1 or the DAKO Envision + System (DAKO) for the others. Signal from the HGF antibody was amplified using biotin-tyramide signal amplification followed by a streptavidin-horseradish peroxidase conjugate (TSA kit; Perkin-Elmer Life Sciences, Boston, MA). Staining was visualized using diaminobenzidine and counterstained with acidified hematoxylin. Slides were also stained in the absence of primary antibody to evaluate nonspecific secondary antibody reactions.

Evaluation of Immunostaining. Immunostaining was scored on a scale of 0 to 3+ (negative/weak/moderate/intense staining). Distinctions between membrane and cytoplasmic staining were impractical given the diffuse staining of the antigens (visualized using the chromogenic substrate, diaminobenzidine). Therefore, scores represent the combined staining intensity of membranous and cytoplasmic staining. Histospots with <10% of their area covered by tumor were excluded from analysis. Scoring was performed by two independent observers (J. Y. K. and M. D.-F.), and histocores with discrepant scores were re-examined by both observers to achieve a consensus score. Cases with scores of 2+ or 3+ were designated as "high," whereas cases with scores of 0 or 1+ were designated as "low."

Statistical Analysis. All of the analyses were completed using Statview 5.0.1 (SAS Institute Inc., Cary, NC). Correlations between markers were performed using a χ² test. Prognostic significance was assessed using both univariate and multivariate Cox proportional hazards models with 30-year survival as an end point. Survival curves were calculated using the Kaplan-Meier method, with significance evaluated using the Mantel-Cox log rank test.

RESULTS

We analyzed the expression of several components of the Met pathway including HGF, matriptase, HAI-1, and Met itself (using both intracellular- and extracellular-specific antibodies; Fig. 1). The number of tumors expressing high levels of these antigens ranged from 18 to 46% (Table 1). Although tissue microarrays have been shown to adequately represent tumor antigen expression (27), they most likely do not fully represent the expression of stromal markers. Therefore, we limited our analysis to the expression of markers by tumor cells and not stroma. In the case of HGF, this resulted in our selective study of autocrine (tumor) expression.

Fig. 1. Immunohistochemical staining of breast cancer tissue microarrays for four members of the HepG2/Met pathway. Each histospot shows a representative positive case (3+) for matriptase (A), HAI-1 (B), HGF (C), and Met (D), at low (×200) and high (×200, insets) magnification. Matriptase, HAI-1, and Met show a membrano-cytoplasmic localization, whereas HGF is predominantly cytoplasmic. The cellular localization of Met using antibodies to the intracellular and extracellular domains was comparable (D, left and right insets, respectively). Representative histospots stained using an antibody to the intracellular domain of Met and scored as 0, 1, 2, or 3+ are shown in E-H, respectively.
levels of HAI-1 (59% *versus* 40% survival; *P* = 0.0110; Table 3). Of the traditional markers of tumor aggression, only tumor size was predictive of outcome (*P* = 0.0001; Table 3). Kaplan-Meier curves demonstrated that these markers were prognostic over the entire 30-year follow-up period (Fig. 2).

Previous studies have suggested that the ratio of HAI-1 and matriptase may play an important role in promoting tumor aggression (13). Therefore, we compared the survival of patients with tumors expressing high and/or low levels of each. Among tumors expressing high levels of matriptase, HAI-1 coexpression predicted a worse outcome (relative risk = 1.88; 95% CI, 1.05–3.36; *P* = 0.0335). Comparison of tumors expressing both markers to those expressing neither demonstrated that patients with double-positive tumors have an increase relative risk of 2.43 (95% CI, 1.36–4.34; *P* = 0.0026). Addition of Met (cytoplasmic domain) to this analysis showed that patients with Met, matriptase, and HAI-1-positive tumors exhibit an increased relative risk of 3.25 (95% CI, 1.24–8.50; *P* = 0.0165).

We then determined the independent predictive power of the Met pathway. First, we limited our analysis solely to the Met pathway components. Using multivariate analysis, both the cytoplasmic tail of Met and HAI-1 retained independent predictive power (Table 4). When we included these two markers with traditional breast cancer markers, they retained their independence. Tumor size was the only other independent predictor of poor outcome (Table 5).

To elucidate potential associations between these markers, we performed *χ²* analyses, which revealed significant associations between the expression of Met, HGF, and matriptase (*P* < 0.0002; Table 2). HAI-1 expression was independent of these markers, although it trended toward coexpression with matriptase (*P* = 0.0597). Interestingly, the expression of antibodies to the intracellular and extracellular domains of Met were highly correlated (*P* < 0.0001; Table 2), but not coincident; 38 cases were scored as entirely Met-extracellular negative (score of 0) and Met-intracellular high (12.2%), whereas only 1 case was judged as entirely Met-intracellular negative and Met-extracellular high (0.3%). This result suggests that relative overexpression of the intracellular domain of Met is far more common than relative overexpression of the extracellular domain.

To determine the predictive power of the Met pathway, we initially performed a univariate analysis of individual Met pathway components and compared them with traditional breast cancer markers (Table 3). Because breast carcinoma can recur and kill patients decades after its initial diagnosis, we studied 30-year disease-related survival. Using univariate analysis, only the cytoplasmic tail of Met, not the extracellular portion, showed prognostic power (*P* = 0.0029; Table 3). Sixty-one percent of patients overexpressing the cytoplasmic domain of Met died of breast cancer within 30 years compared with 41% with lower levels. High-level matriptase expression was also predictive of poor outcome (51% dead of disease *versus* 40%; *P* = 0.0279; Table 3), as were elevated

---

**Table 1 Marker expression**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Low (%)</th>
<th>High (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met (cytoplasmic)</td>
<td>229 (72)</td>
<td>91 (28)</td>
<td>320</td>
</tr>
<tr>
<td>Met (extracellular)</td>
<td>251 (78)</td>
<td>72 (22)</td>
<td>323</td>
</tr>
<tr>
<td>Matriptase</td>
<td>181 (55)</td>
<td>148 (45)</td>
<td>329</td>
</tr>
<tr>
<td>HGF</td>
<td>176 (54)</td>
<td>147 (46)</td>
<td>323</td>
</tr>
<tr>
<td>HAI-1</td>
<td>260 (82)</td>
<td>56 (18)</td>
<td>316</td>
</tr>
<tr>
<td>ER</td>
<td>212 (69)</td>
<td>39 (13)</td>
<td>351</td>
</tr>
<tr>
<td>PR</td>
<td>203 (68)</td>
<td>57 (19)</td>
<td>307</td>
</tr>
<tr>
<td>HER2 neu</td>
<td>260 (87)</td>
<td>56 (19)</td>
<td>297</td>
</tr>
<tr>
<td>Nuclear grade III</td>
<td>241 (81)</td>
<td>56 (19)</td>
<td>297</td>
</tr>
<tr>
<td>Patient age (&gt;50 yr)</td>
<td>225 (68)</td>
<td>56 (19)</td>
<td>307</td>
</tr>
<tr>
<td>Tumor size (&gt;2 cm)</td>
<td>225 (68)</td>
<td>56 (19)</td>
<td>307</td>
</tr>
</tbody>
</table>

---

**Table 2 Met pathway associations: *χ²* analysis**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Met (cyto)</th>
<th>Met (extra)</th>
<th>Matriptase</th>
<th>HGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met (extra)</td>
<td>&lt;0.0001</td>
<td>0.288</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matriptase</td>
<td>&lt;0.0001</td>
<td>0.0002</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>HGF</td>
<td>0.3716</td>
<td>0.8330</td>
<td>0.0597</td>
<td>0.2902</td>
</tr>
</tbody>
</table>

---

**Table 3 Univariate analysis**

<table>
<thead>
<tr>
<th>Marker—high expression</th>
<th><em>P</em></th>
<th>Relative risk</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met (cytoplasmic)</td>
<td>0.0029</td>
<td>1.826</td>
<td>1.228–2.715</td>
</tr>
<tr>
<td>Met (extracellular)</td>
<td>0.0771</td>
<td>1.007</td>
<td>0.639–1.585</td>
</tr>
<tr>
<td>Matriptase</td>
<td>0.0279</td>
<td>1.527</td>
<td>1.047–2.169</td>
</tr>
<tr>
<td>HGF</td>
<td>0.5026</td>
<td>1.140</td>
<td>0.778–1.670</td>
</tr>
<tr>
<td>HAI-1</td>
<td>0.0110</td>
<td>1.808</td>
<td>1.145–2.533</td>
</tr>
<tr>
<td>BR</td>
<td>0.3663</td>
<td>1.214</td>
<td>0.797–1.850</td>
</tr>
<tr>
<td>PR</td>
<td>0.7881</td>
<td>0.945</td>
<td>0.627–1.424</td>
</tr>
<tr>
<td>HER2/neu</td>
<td>0.4535</td>
<td>0.786</td>
<td>0.420–1.473</td>
</tr>
<tr>
<td>Nuclear grade III</td>
<td>0.7993</td>
<td>0.934</td>
<td>0.555–1.574</td>
</tr>
<tr>
<td>Patient age (&gt;50 yr)</td>
<td>0.3240</td>
<td>1.227</td>
<td>0.817–1.843</td>
</tr>
<tr>
<td>Tumor size (&gt;2 cm)</td>
<td>0.0001</td>
<td>2.134</td>
<td>1.448–3.145</td>
</tr>
</tbody>
</table>

---

**Table 4 Multivariate analysis: Met pathway**

<table>
<thead>
<tr>
<th>Marker—high expression</th>
<th><em>P</em></th>
<th>Relative risk</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met (cytoplasmic)</td>
<td>0.0064</td>
<td>1.862</td>
<td>1.191–2.910</td>
</tr>
<tr>
<td>Met (extracellular)</td>
<td>0.7710</td>
<td>1.079</td>
<td>0.645–1.805</td>
</tr>
<tr>
<td>Matriptase</td>
<td>0.2413</td>
<td>1.290</td>
<td>0.843–1.974</td>
</tr>
<tr>
<td>HGF</td>
<td>0.5189</td>
<td>1.159</td>
<td>0.740–1.815</td>
</tr>
<tr>
<td>HAI-1</td>
<td>0.0291</td>
<td>1.721</td>
<td>1.057–2.803</td>
</tr>
</tbody>
</table>

---

* Multivariate analysis of 30-year disease-related survival was performed using the Cox-proportional hazards model. Statistically significant observations are in boldface.

Fig. 2. Kaplan-Meier analysis of disease-related survival demonstrates that Met (intracellular domain), matriptase, and HAI-1 show long-term prognostic benefit. Statistical significance was assessed using the log rank test.

---

1103
Table 5  Multivariate analysis: traditional markers

<table>
<thead>
<tr>
<th>Marker—high expression</th>
<th>P</th>
<th>Relative risk</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met (cytoplasmic)</td>
<td>0.0098</td>
<td>1.841</td>
<td>1.159–2.925</td>
</tr>
<tr>
<td>HAI-1</td>
<td>0.0483</td>
<td>1.805</td>
<td>1.003–3.242</td>
</tr>
<tr>
<td>ER</td>
<td>0.8552</td>
<td>1.597</td>
<td>0.568–4.197</td>
</tr>
<tr>
<td>PR</td>
<td>0.9040</td>
<td>0.970</td>
<td>0.590–1.596</td>
</tr>
<tr>
<td>HER2/neu</td>
<td>0.5284</td>
<td>0.783</td>
<td>0.367–1.674</td>
</tr>
<tr>
<td>Nuclear grade III</td>
<td>0.7886</td>
<td>1.084</td>
<td>0.602–1.952</td>
</tr>
<tr>
<td>Patient age (&gt;50 yr)</td>
<td>0.5320</td>
<td>1.177</td>
<td>0.706–1.963</td>
</tr>
<tr>
<td>Tumor size (&gt;2 cm)</td>
<td>0.0007</td>
<td>2.265</td>
<td>1.415–3.627</td>
</tr>
</tbody>
</table>

Multivariate analysis of 30-year disease-related survival was performed using the Cox-proportional hazards model using both the independent markers from Table 4 and traditional histopathologic measures. Statistically significant observations are in boldface.

**DISCUSSION**

The relative importance of ligand-dependent and ligand-independ-ent Met activation in carcinogenesis is a matter of continued debate (1). Our studies demonstrate that members of the HGF pathway, namely HGF, Met, and matriptase, are often coexpressed on breast cancers, and that high-level expression of two of these members, Met and matriptase, associates with more aggressive tumors. Such observations would be expected if HGF-mediated Met stimulation played a role in tumorigenesis. Although tumoral HGF levels were not found to be predictive of outcome, the expression of matriptase was prognostic in an univariate analysis, suggesting that as an activator of HGF, it plays a rate-dependent ("gate-keeping") role in the ligand-dependent stimulation of Met.

In addition to their role in the activation of HGF, matriptase and its cognate inhibitor, HAI-1, also play a role in the plasminogen activator cascade (6, 31, 32). This cascade culminates in the activation of plasmin and the coactivation of matrix metalloproteinases, both of which degrade extracellular matrix components and potentiate tumor cell invasion, extravasation, and metastasis (33). Matriptase promotes this pathway by activating latent uPA, which, in turn, activates plasmin (6, 34). Like matriptase, uPA can also cleave pro-HGF, providing another level of interaction between the Met/HGF pathway and plasmin cascade (35). Given the multiple functions of matriptase, it is not surprising that aggressive breast tumors produce higher levels of this enzyme.

Likewise, HAI-1, as an inhibitor of matriptase, may help modulate both the Met/HGF and plasmin pathways. Interestingly, the expression of HAI-1 was independent of the other members of the Met pathway indicating that its expression is regulated differently. Previous reports have suggested that HAI-1 is down-regulated in colon carcinoma and high-grade ovarian carcinomas (13, 36). In contrast, our study demonstrates that HAI-1 expression is associated with aggressive breast carcinomas, being an independent predictor of poor outcome (Table 5). Although this may be puzzling in light of the role of HAI-1 in inhibiting HGF-dependent Met activation, the coordinated expression of both matriptase and HAI-1 may be far more important in promoting tumor aggression than the unopposed production of active matriptase in the absence of its inhibitor. Coordinated regulation of another inhibitor of the uPA/plasmin cascade, PAI-1, is crucial for inducing tumor invasion (37). Furthermore, a recent meta-analysis of the uPA/PAI-1 system in breast cancer demonstrated conclusively that overexpression of both uPA and its inhibitor PAI-1 were associated with poor outcome in breast cancer (32). In our study, the importance of matriptase and HAI-1 coexpression is demonstrated in the elevated relative risk of patients with tumor expressing both markers.

Because of lot-to-lot inconsistencies in polyclonal Met-antibodies, we analyzed two different monoclonal antibodies, one to the extra-

**ACKNOWLEDGMENTS**

We thank Thomas D’Aquila and Lori Charette for their help in this effort.

**REFERENCES**


Tissue Microarray-Based Studies of Patients with Lymph Node Negative Breast Carcinoma Show that Met Expression Is Associated with Worse Outcome but Is Not Correlated with Epidermal Growth Factor Family Receptors

Idris Tolgay Ocal, M.D.
Marisa Dolled-Filhart, B.S.
Thomas G. D'Aquila, B.S.
Robert L. Camp, M.D., Ph.D.
David L. Rimm, M.D., Ph.D.

Department of Pathology, Yale University School of Medicine, New Haven, Connecticut.

BACKGROUND. It has been shown that receptor tyrosine kinases (RTKs) predict outcome in patients with breast carcinoma. Although RTKs are a large family, HER-2, epidermal growth factor receptor (EGFR), Met (hepatocyte growth factor receptor), and others all have shown the ability to predict outcome. However, it remains unclear whether these markers are defining the same subpopulation of patients with breast carcinoma. In this study, the authors attempted to determine the correlation between RTKs on the basis of their ability to stratify a population according to outcome.

METHODS. The authors used tissue microarray technology to study 324 patients with lymph node negative breast carcinoma who had 20–40 years of follow-up. Expression was assessed using immunohistochemical stains for Met, EGFR, fibroblast growth factor receptor (FGFR), and HER-2. Expression levels were assessed by two observers, and correlations were analyzed. Standard pathology information, including tumor size, nuclear grade, Ki-67 receptor status, and estrogen and progesterone receptor expression levels, also was collected.

RESULTS. RTK expression in the study cohort revealed two strong correlations. Specifically, HER-2 and EGFR showed similar expression patterns (P < 0.0001), and Met cytoplasmic domain and FGFR cytoplasmic staining showed similar expression patterns (P < 0.0001), but no correlation was found between the two groups. Of these RTKs, only high levels of Met cytoplasmic domain showed significance as a prognostic marker defining a shortened survival compared with the rest of the population (P = 0.0035; relative risk, 2.04). In the same group of patients, HER-2, hormone receptor status, and other RTK family receptors were not correlated with outcome. In multivariate analysis, only Met cytoplasmic domain and tumor size showed independent predictive value.

CONCLUSIONS. The current results indicate that the cytoplasmic domain of Met shows a unique staining pattern and defines a set of patients unique from the set of patients defined by overexpression of HER-2, EGFR, or hormone receptors. Furthermore, this group of patients is associated tightly and independently with worse outcome. Cancer 2003;97:1841–8. © 2003 American Cancer Society.
DOI 10.1002/cncr.11335

KEYWORDS: hepatocyte growth factor, c-met, adhesion, scatter factor, prognostic, ErbB2.
The single best indicator of disease free survival and overall survival in patients with breast carcinoma is lymph node status. Patients who have breast carcinomas with axillary lymph node metastases have a 10-year recurrence rate approaching 70%. However, in patients with lymph node negative breast carcinoma, there are no markers used to predict outcome that consistently show statistical significance. In this group, predicting a worse outcome is of great importance, because as many as 20% of women with lymph node negative breast carcinoma eventually will die of metastatic disease. Numerous prognostic markers have been tested to try to predict outcome in this group. In large studies with long follow-up, the best predictors of outcome were tumor size, tumor grade, cathepsin-D expression, Ki-67 expression, S-phase fraction, mitotic index, and vascular invasion. However, despite their marginal statistical status, tumor size and tumor grade enjoy broad acceptance as prognostic factors this group of patients.

Tyrosine kinase receptors (RTKs) are gaining attention as prognostic markers and as possible future predictive markers as the number of trials grows for biospecific inhibitors. Among these, HER-2 (erb B2 and neu), epidermal growth factor receptor (EGFR), and Met (c-met) have been documented as prognostically significant markers for invasive breast carcinomas predicting a worse prognosis. HER-2 has been associated with outcome in patients with lymph node positive tumors but has not proven valuable in patients with lymph node negative tumors. Although some studies have suggested that EGFR is valuable, EGFR has been examined in over 25 studies, of which only approximately 50% suggest that overexpression is associated with poor outcome. Met has also shown mixed results. Although we and others have found Met useful in predicting worse outcome, others either have not seen the correlation or have found an opposite correlation. Finally, FGFR has been studied less and has shown no definitive prognostic value for patients with breast carcinoma.

Slide-to-slide standardization is a long-standing problem in immunohistochemistry studies and may be one explanation for the variability seen in the studies described earlier. Tissue microarray technology can eliminate this problem. Tissue microarrays are a method of placing very small samples of tissue from hundreds or thousands of patients on a single slide. The technology has been used extensively and is the subject of multiple reviews. Tissue microarrays are suited especially well to comparisons of expression between multiple prognostic markers. In this report, we revisit the issue of the prognostic value of four RTKs in patients with lymph node negative breast carcinoma with an emphasis on the correlation between the markers. We studied expression patterns of Met, EGFR, fibroblast growth factor receptor (FGFR), and HER-2/neu on a cohort of patients with long-term follow-up.

MATERIALS AND METHODS

Tissue Microarray Construction

The tissue microarrays were constructed as described previously and as reviewed recently. Briefly, formalin fixed, paraffin embedded tissue blocks containing breast carcinoma specimens were retrieved from the archives of the Yale University Department of Pathology. Areas of invasive carcinoma were identified on corresponding hematoxylin and eosin-stained slides, and the tissue blocks were cored and transferred to a recipient master block using a Tissue Microarrayer (Beecher Instruments, Sun Prairie, WI). Each core measured 0.6 mm in greatest dimension, and cores were spaced 0.8 mm apart. After cutting the recipient block and transfer with an adhesive tape to coated slides for subsequent ultraviolet cross linkage (Instrumedics, Inc., Hackensack, NJ), the slides were dipped in a layer of paraffin to prevent oxidation. The array for the cohort of patients with lymph node negative breast carcinoma was constructed from paraffin embedded, formalin fixed tissue blocks from the Yale University Department of Pathology archives. The specimens were resected between 1962 and 1980, with a follow-up that ranged between 4 months and 53.8 years, with a mean follow-up of 15.6 years and a median follow-up of 14.3 years.

Grading

Nuclear grade was evaluated by one observer (I.T.O.) according to the methods described by Fisher et al. on each spot. Due to the age of the specimens, they were not assigned a nuclear grade previously. Presence or absence of necrosis, mitotic count, or histologic parameters could not be included in the grading criteria because of the small size of the area evaluated. Eighty of 306 included specimens (26%) were Grade 1, 170 specimens (56%) were Grade 2, and 56 specimens (18%) were Grade 3. Similar to previous studies of nuclear grade in lymph node negative tumors, no statistical significance was achieved. However, a trend was seen between high nuclear grade and worse outcome (5-year survival).

Tumor Size

Information regarding tumor size was obtained from descriptions in the original pathology reports. For staging in the statistical analyses, tumor size > 2.0 cm
was considered *large*, and others were considered *small*.

**Immunohistochemistry**
The tissue microarray slides were deparaffinized with xylene rinses and then transferred through two changes of 100% ethanol. Endogenous peroxidase activity was blocked by a 30-minute incubation in a 2.5% hydrogen peroxide/methanol buffer. Antigen retrieval was performed by boiling the slides in a pressure cooker filled with a sodium citrate buffer, pH 6.0. After antigen retrieval, the slides were incubated with 0.3% bovine serum albumin/1 × Tris-buffered saline (TBS) for 1 hour at room temperature to reduce nonspecific background staining, followed by a series of 2-minute rinses in 1 × TBS, TBS/0.01% Triton, and 1 × TBS. Primary antibody was applied for 1 hour at room temperature. Dilutions for the RTKs were as follows: Met, 1:1 (note that this antibody was provided as a culture supernate from Zymed Laboratories, Inc., South San Francisco, CA); EGFR, 1:200; FGFR, 1:300; and HER-2, 1:8000. After a series of TBS rinses, as described above, bound antibody was detected by using an antirabbit, horseradish peroxidase-labeled, polymer secondary antibody from the DAKO Envision TM + System (DAKO, Carpinteria, CA). The slides were rinsed in the TBS series and visualized with a 10-minute incubation of liquid 3,3′-diaminobenzidine in buffered substrate (DAKO) for 10 minutes. Finally, the slides were counterstained with hematoxylin and mounted with Immunomount (Shandon, Pittsburgh, PA). Immunohistochemical staining also was done for estrogen receptor (ER), progesterone receptor (PR), and HER-2, as described previously. 26 Ki-67 expression was assessed using purified antihuman monoclonal antibody (1:200 dilution; overnight incubation; Pharmingen, San Diego, CA). The Met antibody 3D4 was provided as part of a collaborative arrangement with Zymed Laboratories, Inc.; the EGFR antibody was EGFR 1005 (catalog no. SC03) from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA); and the FGFR1 antibody was FLG C15 (catalog no. SC121; Santa Cruz Biotechnology, Inc.).

**Evaluation of Immunohistochemical Staining**
For each spot, the regions of most intense and/or predominant staining pattern were scored by eye. Traditionally, immunohistochemistry scoring of stain intensity includes a variable for the area percentage stained with the specimen; however, due to the small size of the spot (0.6 mm in greatest dimension) and the fact that the spots often are homogenous, no area variable was included. Nuclear staining and/or cytoplasmic staining (ER/PR, EGFR, FGFR, Ki-67, and Met cytoplasmic domains) were determined separately for each specimen. The staining intensity was graded on the following scale: 0, no staining; 1, weak staining; 2, moderate staining; and 3, intense staining. The membranous staining was determined for HER-2 and EGFR. The staining intensity for membranous staining was graded on the following scale: 0, no staining; 1, incomplete staining; 2, weak but complete staining of the plasma membrane encircling the entire cell; and 3, intense complete staining. Again, we did not take the percentage of cells with staining in consideration because of the small size of the tumor sections. For specimens that were uninterpretable, a score of not available was given. Only FGFR showed distinctly separate nuclear and cytoplasmic staining characteristics; and, for this antibody, nuclear and cytoplasmic staining were scored individually. Scoring of the tissue microarrays was completed by two independent observers (I.T.O and M.D.F.) for EGFR, FGFR nuclear staining, FGFR cytoplasmic staining, and Met with a very high correlation between scorers (P < 0.0001). For the antibodies with established staining characteristics in the literature (ER, PR, HER-2, and Ki-67), scoring was performed by one observer (I.T.O. or M.D.F.). Ki-67 was considered positive if > 10% of nuclei were stained. Frequency distributions for these markers were in the range of other works seen in the literature, validating this cohort (see Table 1).

**Statistical Analysis**
All analyses were completed using Statview software (version 5.0.1; SAS Institute Inc., Cary, NC). The correlation between the scores of both scorers and the relations among the different immunohistochemical and clinicopathologic parameters were measured using the chi-square test. The prognostic significance of parameters on overall survival was calculated by multivariate analysis using a Cox proportional hazards model. Survival curves were calculated using the Kaplan–Meier survival analysis method with the differences estimated using the Mantel–Cox log-rank test.
RESULTS

The RTKs analyzed in this study showed a variety of staining patterns that are summarized in Figure 1. Although all RTKs are present at the membrane, there are now numerous studies showing that RTKs can be found in other locations in the cell, including both the cytoplasm and, more recently, the nucleus.\(^2\) The patterns we saw included both conventional membranous patterns as well as both cytoplasmic and nuclear staining (for a detailed description, see above). The pattern of expression of each RTK was determined after examining the entire array to determine the most prevalent patterns. The question of how to divide subjective ordinal staining patterns always is controver-
sial. In this study, we tried to use breakpoints suggested previously in the literature when they were present\(^8,10\); however, in some instances (FGFR), there were not clear precedents in the literature or the literature was inconsistent (EGFR). A summary of the expression pattern of each RTK is shown in Table 2 with the definition of the cut-off values.

In this study, we were concerned primarily with the correlation between each RTK and the correlation of each RTK with known prognostic markers. We calculated chi-square \(P\) values to determine the correlation among expression patterns of individual parameters. Table 3 shows the chi-square \(P\) values for all parameters studied in patients with lymph node neg-
active breast carcinoma. Of particular note are the highly significant correlations in immunohistochemical expression patterns of HER-2 and EGFR and of Met and FGFR cytoplasmic staining ($P < 0.0001$). ER expression was correlated highly with PR expression, and they both showed a significant, inverse correlation with HER-2 expression.

Survival Analyses
Cox univariate analyses at 10 years for all variables studied in patients with lymph node negative breast carcinoma are shown in Table 4. Ki-67 and tumor size, as expected, were correlated with poorer survival (Ki-67: $P = 0.04$; relative risk [RR], 1.648; tumor size: $P = 0.0008$; RR, 2.201). Among the other variables studied, only the cytoplasmic domain of Met showed a statistically significant correlation with a worse prognosis and shortened survival ($P = 0.0035$; RR, 2.041). Survival curves were produced for each variable, but only Met reached statistical significance using the Mantel–Cox log-rank test (Fig. 2). To determine the independent predictive value of Met expression, a multivariate analysis was conducted using the Cox proportional hazards model. In multivariate analysis, Met retained its significance as a predictor of worse outcome, even when the model contained all of the conventional prognostic variables as well as all other RTKs tested (RR, 1.86; $P = 0.011$).

DISCUSSION
In patients with lymph node negative breast carcinoma, we still are unable to discern the 15–20% of patients who eventually will succumb to their disease. A long list of potential molecular markers of poorer prognosis have been suggested for this group patients with breast carcinoma, including Ki-67, cathepsin-D, HER-2, p53, low levels of nm23, and hormone receptors. There also are reports that histologic findings (e.g., tumor grade, tumor size, mitotic index, and vascular invasion) and cytometric data (e.g., DNA ploidy and S-phase fraction) are helpful. To date, there are no universally accepted markers for this group of patients.

We evaluated our tissue microarray cohort using these traditional markers for breast carcinoma. Hormone receptors showed similar staining patterns and frequencies compared with the widely reported values in breast carcinoma. ER was negative in 35% of patients, and PR was negative in 40% of patients. Neither ER expression nor PR expression was predictive of survival in our group. This finding is consistent with many other published studies showing limited value or no value for ER and PR as prognostic markers in patients with lymph node negative breast carcinoma. HER-2 has been proven as a predictor of worse prognosis in patients with invasive, lymph node positive breast lesions and is a better predictor of survival compared with hormone receptors for this group of patients. However, its prognostic value in patients with lymph node negative lesions remains controversial, with conflicting data in the literature. Our findings appear to support the hypothesis that HER-2 is not a reliable prognostic marker in this group of patients.

EGFR expression in breast carcinomas has been reported, both in patients with lymph node negative breast carcinoma and in patients with lymph node positive breast carcinoma, as associated with poor prognosis; however, the clinical significance and association with disease free survival (DFS) and overall survival statistics show mixed results in different studies. Tsutui et al. reported that EGFR carries a prognostic significance only for DFS in patients with lymph node negative tumors on multivariate analysis; whereas, in the series by Torregrosa et al., EGFR failed to show statistical significance as a prognostic marker in patients with lymph node negative tumors but was associated with a worse prognosis, influencing DFS in patients with lymph node positive tumors. Conversely, Seshadri et al. concluded that the expression of EGFR was not at all a predictor of poor prognosis in patients with lymph node negative breast carcinoma. Our findings did not show any value for EGFR as a marker of prognosis in patients with lymph node negative tumors.

Our group has been particularly interested in the hepatocyte growth factor receptor, Met (or C-met). Met is a dimeric tyrosine kinase growth factor receptor in which activation is associated with increased invasion, motogenesis, and morphogenesis. In the breast, Met is expressed in normal ductal and lobular
TABLE 3
Chi-Square Analysis of the Correlation Between Expression Levels of Each Marker

<table>
<thead>
<tr>
<th>Marker</th>
<th>ER</th>
<th>PR</th>
<th>HER-2</th>
<th>EGFR</th>
<th>FGFR-c</th>
<th>FGFR-n</th>
<th>MET</th>
<th>KI-67</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>&lt; 0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR</td>
<td>0.0003&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>0.0011&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HER-2</td>
<td>&lt; 0.0001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 0.0001&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>0.17</td>
<td>0.17</td>
<td>0.074</td>
<td>0.11</td>
<td>0.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGFR-c</td>
<td>0.55</td>
<td>0.029&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>0.038&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.033</td>
<td>0.007&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.26</td>
<td>0.03&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>FGFR-n</td>
<td>0.01&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>0.017&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>0.019&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.0475&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.065</td>
<td>&lt; 0.0001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 0.0001&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>MET</td>
<td>0.03&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.19</td>
<td>0.54</td>
<td>0.77</td>
<td>&lt; 0.0001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KI-67</td>
<td>0.55</td>
<td>0.029&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>0.038&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.033</td>
<td>0.007&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.26</td>
<td>0.03&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

ER: estrogen receptor; PR: progesterone receptor; EGFR: epidermal growth factor receptor; FGFR-c: fibroblast growth factor receptor (FGFR) cytoplasmic staining; FGFR-n: FGFR nuclear staining.

<sup>a</sup>Statistically significant correlation patterns among the parameters studied in patients with lymph node-negative breast carcinoma (n = 364 patients).

<sup>b</sup>Statistically significant.

<sup>c</sup>Inverse correlation.

TABLE 4
Univariate Analysis of Conventional and Receptor Tyrosine Kinase Markers

<table>
<thead>
<tr>
<th>Marker</th>
<th>RR</th>
<th>P value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>1.031</td>
<td>0.9042</td>
<td>0.633-1.681</td>
</tr>
<tr>
<td>PR</td>
<td>0.946</td>
<td>0.8184</td>
<td>0.587-1.522</td>
</tr>
<tr>
<td>HER-2</td>
<td>1.118</td>
<td>0.7236</td>
<td>0.603-2.071</td>
</tr>
<tr>
<td>EGFR</td>
<td>1.029</td>
<td>0.9400</td>
<td>0.494-2.141</td>
</tr>
<tr>
<td>FGFR-c</td>
<td>1.078</td>
<td>0.7456</td>
<td>0.685-1.695</td>
</tr>
<tr>
<td>FGFR-n</td>
<td>1.338</td>
<td>0.2200</td>
<td>0.840-2.129</td>
</tr>
<tr>
<td>MET</td>
<td>2.041&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.0035&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.264-3.295&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>KI-67</td>
<td>1.648&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.0148&lt;sup&gt;*&lt;/sup&gt;</td>
<td>1.019-2.665&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Nuclear grade</td>
<td>1.033</td>
<td>0.9120</td>
<td>0.578-1.846</td>
</tr>
<tr>
<td>Tumor size</td>
<td>2.201&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.0006&lt;sup&gt;*&lt;/sup&gt;</td>
<td>1.391-3.401&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

95% CI: 95% confidence interval; ER: estrogen receptor; PR: progesterone receptor; EGFR: epidermal growth factor receptor; FGFR-c: fibroblast growth factor receptor (FGFR) cytoplasmic staining; FGFR-n: FGFR nuclear staining.

<sup>a</sup>Statistically significant.

epithelium and functions in both the embryonic development and subsequent remodeling of the breast.42

Our previous studies showed that expression of the Met receptor in patients with invasive breast carcinoma is of significant prognostic value in determining patient survival, even in patients with negative lymph nodes.6,9 The Vande Woude group also found this correlation between Met expression and outcome;<sup>10</sup> although, along with others, we found no correlation in some cohorts (unpublished data). We believe this variability is due to antibody selection. Many studies have used antibodies to the c-terminal domain produced by Santa Cruz Biotechnology, Inc., using a peptide coding for the C-terminal 28 amino acids. We found significant lot-to-lot variability with this antibody. Recently, Zymed Laboratories, Inc. and others have produced monoclonal antibodies to the C-terminus. In another study from our laboratory,<sup>43</sup> we compared the results from this antibody with another monoclonal antibody to the extracellular domain of Met. We showed results with the cytoplasmic domain similar to the results reported here, but the extracellular domain was very different. Although there is a high correlation of expression, overexpression, as assessed by the antibody to the cytoplasmic domain, selects a group of patients with worse outcome, whereas the extracellular domain antibody does not.<sup>43</sup> We believe this may be a function of either cleavage or activation of Met. It is notable that the other study that found a correlation between Met expression and patient outcome also used a monoclonal antibody (generated by the Vande Woude laborato-
ry) and that our own previous studies used a polyclonal antibody made to a cytoplasmic domain peptide.

The finding that Met overexpression predicts poor outcome raises the question of its relation to other RTKs, especially EGFR and HER-2, which also have been implicated as prognostic variables. The current work shows that tumors that overexpress Met are unique from tumors that express HER-2 and EGFR. EGFR and HER-2 are closely related members of the erbB oncogene family; thus, it is not surprising that there is a high correlation between the expression of these two proteins. Met is not a member of this family; thus, its overexpression appears to be unrelated to erbB family RTKs.

Although there is not a close correlation between Met and the erbB family of RTKs, there is a tight, direct correlation between Met expression and the cytoplasmic staining pattern seen with the FGFR antibody (P < 0.0001). To our knowledge, the correlation between FGFR and Met has never been reported before in patients with breast carcinoma. Coordinated actions of several growth factors and their receptors, including FGFR and C-met, have been reported in patients with hepatic lesions and in normal morphogenesis in the uterus; however, the molecular basis of such an interaction, if present, remains unknown. Although FGFR is a membrane-based RTK, immunohistochemical evaluation for FGFR has not been standardized as a marker for immunohistochemistry. Thus, we scored cytoplasmic and nuclear staining for FGFR separately. Although neither cytoplasmic staining for FGFR nor nuclear staining for FGFR was a statistically significant prognostic marker in our study, there was a high correlation between Met expression and FGFR cytoplasmic expression.

RTKs recently have gained great interest both as markers of prognosis and as promising targets for novel chemotherapy options. The HER2/trastuzumab pair is the first U.S. Food and Drug Administration-approved example of numerous RTK-related therapeutics currently in clinical trials. This interest raises the question of the relations between different RTKs. Although data on the cross reactivity of kinase-based therapeutics is not widely available to date, there are many types of RTK receptors that are overexpressed in malignancies of the breast.

The results of the current study indicate that the expression of Met, as assessed using a cytoplasmic domain monoclonal antibody, in patients with lymph node negative, invasive breast carcinoma is of significant predictive value in determining patient survival. Its predictive value exceeds all conventional prognostic markers assessed in this cohort as well as the other RTKs tested. The group of patients that overexpress this protein are unrelated to the group that overexpresses either EGFR or HER-2. Thus, in the future, determination of prognosis in patients with lymph node negative, invasive breast carcinoma may be improved by assessment of the level of expression of Met using a cytoplasmic domain monoclonal antibody.

REFERENCES


32. Rosen PR, Groshen S, Saigo PE, Kinne DW, Hellman S. A long-term follow-up study of survival in Stage I (T1N0M0) and Stage II (T1N1M0) breast carcinoma. *J Clin Oncol*. 1989; 7:355–366.


Quantitative Analysis of Breast Cancer Tissue Microarrays Shows That Both High and Normal Levels of HER2 Expression Are Associated with Poor Outcome

Robert L. Camp, Marisa Dolsed-Filhart, Bonnie L. King, and David L. Rimm

Departments of Pathology [R. L. C., D. L. R.], Genetics [M. D.-F.], and Therapeutic Radiology [B. L. K.], Yale University School of Medicine, New Haven, Connecticut 06520

Abstract

Using a tissue microarray cohort of 300 breast cancers and 84 samples of normal breast epithelium, we analyzed HER2/neu expression and compared traditional clinical (manual) scoring with a recently developed system for the quantitative measurement of immunohistochemical stains (AQA). As expected, both methods identified a population (10–15%) of high-HER2-expressing tumors with poor 30-year disease-related survival. Using AQA analysis, we found that normal epithelium expresses a low but detectable level of HER2 and that 17.5% of tumors exhibit similar low-level HER2 expression. This low group was not definable by manual scoring. Surprisingly, HER2-normal tumors were as aggressive as HER2-overexpressing tumors. Our study suggests that in situ quantitative measurement of HER2 stratifies breast tumors into three expression levels: normal, intermediate, and high, where both normal and high levels are associated with a worse outcome.

Introduction

HER2 (neu or erb-B2), a member of the epidermal growth factor family, is genetically amplified and overexpressed in aggressive breast cancers. High levels of HER2 are associated with poor prognosis, particularly in node-positive breast carcinoma patients. Recently, a targeted therapeutic agent against HER2 has been developed. Trastuzumab (Herceptin) is a humanized monoclonal antibody directed against the extracellular domain of HER2. Treatment of patients with metastatic breast carcinoma with Herceptin has shown therapeutic benefit, especially when combined with conventional chemotherapy agents. The association between HER2 expression and Herceptin response has stimulated renewed interest in accurately assessing HER2 amplification and overexpression. Toward this goal, we have developed a system for compartmentalized, automated quantitative analysis of histological sections (AQA; Ref. 1). As with an ELISA, AQA provides highly reproducible analysis of target signal expression with use of a continuous, rather than nominal, scale. Unlike an ELISA, spatial information, including tissue and subcellular localization, is preserved. Using a tissue microarray composed of archival breast cancer specimens and normal epithelia, we found a bimodal distribution of HER2, where tumors expressing both high and normal HER2 levels exhibited poor 30-year disease-specific survival.

Materials and Methods

Tissue Microarray Design. Paraffin-embedded, formalin-fixed specimens from 300 cases of node-positive invasive breast carcinoma were identified from the archives of the Yale University Department of Pathology as available from 1959 to 1977, with a mean follow-up time of 9.6 years. No patients received Herceptin during the study period. Complete treatment information was unavailable for the entire cohort; however, most patients were treated with local radiation and ~15% were treated with chemotherapy consisting primarily of Adriamycin, cytoxan, and 5-fluorouracil. Approximately 27% subsequently received tamoxifen (post-1978). Seven patients had biopsy-proven stage IV disease at the time of diagnosis.

In constructing the microarrays, we identified areas of invasive carcinoma, away from in situ lesions and normal epithelium, and took two 0.6-mm cores. We cut 5-μm-thick sections of the microarrays and processed them as described previously (2, 3). We previously demonstrated with HER2 that two cores replicated the results of an entire slide in >95% of cases (4). An additional microarray consisting of 84 samples of normal epithelium was also constructed from samples of normal ducts and lobules taken from breast cancer patients. Samples were taken away from areas of tumor and assessed histologically to ensure that they were not affected by atypical hyperplasia or carcinoma in situ.

Immunohistochemistry. Tissue microarray slides were stained as described (1). In brief, for both manual and automated analysis, slides were incubated for 1 h at room temperature with polyclonal anti-HER2 (1:2000; DAKO Corp., Carpinteria, CA) diluted in Tris-buffered saline containing BSA. Previous analysis of titrations of the HER2 antibody demonstrated that higher dilutions of anti-HER2 antibody (1:1000–1:8000) more accurately define the HER2-high from the HER2-intermediate populations, whereas lower dilutions (1:50–1:500) distinguish the HER2-normal from HER2-intermediate populations.2 In this study we used a concentration (1:200) that sufficiently distinguished all three populations. Goat antirabbit antibody conjugated to horse-radish peroxidase–decorated dextran polymer backbone (Envision; DAKO Corp.) was used as a secondary reagent. For manual analysis, slides were visualized with diaminobenzidine (Dako Corp.), followed by ammonium hydroxide–acidified hematoxylin. For automated analysis, tumor cells were identified by use of a fluorescently tagged anticytokeratin antibody cocktail (AE1/AE3; DAKO Corp.). We added 4',6-diamidino-2-phenylindole to visualize nuclei, and HER2 was visualized with a fluorescent chromogen (Cy-5; tyramide; NEN Life Science Products, Boston, MA). Cy-5 (red) was used because its emission peak is well outside the green-orange spectrum of tissue autofluorescence.

Automated Image Acquisition and Analysis. Automated image acquisition and analysis using AQA has been described previously (1). In brief, monochromatic, high-resolution (1024 × 1024 pixel; 0.5-μm) images were obtained of each histoslot. We distinguished areas of tumor from stromal elements by creating a mask from the cytokeratin signal. Coalescence of cytokeratin at the cell surface helped localize the cell membranes, and 4',6-diamidino-2-phenylindole was used to identify nuclei. The HER2 signal from the membrane area of tumor cells was scored on a scale of 0–255 and expressed as signal intensity divided by the membrane area.

FISH. FISH3 analysis was performed with the PathVysion HER2 DNA Probe Kit (Vysis, Downers Grove, IL), using two directly labeled fluorescent DNA probes complementary to the HER2/neu gene locus (LSI HER2/neu SpectrumRed) and to chromosome 17 pericentromeric α satellite DNA (CEP17 SpectrumGreen), according to standard protocols. HER2/neu gene amplification was quantified by comparing the ratio of LSI HER2/neu to CEP17 probe signals in accordance with the PathVysion HER2 DNA Probe Kit criteria. We examined 60 nonoverlapping tumor cell nuclei in each histo-
spot to determine the average number of HER2/neu and chromosome 17 copies/cell for each tissue specimen. The ratio of these averages was used to determine the presence of HER2/neu gene amplification. Specimens with a HER2/neu:chromosome 17 ratio $>2$ were scored as positive for HER2/neu gene amplification.

Data Analysis. Manual scoring of HER2 expression was assessed by a pathologist (R. L. C.) using a nominal four-point scale (0 to 3+). Histospots containing <10% tumor, as assessed either subjectively (manual) or by mask area (automated), were excluded from further analysis. Previous studies have demonstrated that the staining from a single histospot provides a sufficiently representative sample for analysis (4, 5). Correlations with other prognostic markers were determined by $\chi^2$ analysis. Overall survival analysis was assessed by Kaplan-Meier analysis with the Mantel-Cox log-rank score for determining statistical significance. Relative risk was assessed by the univariate and multivariate Cox proportional hazards model. Analyses were performed with Statview 5.0.1 (SAS Institute, Cary, NC). Patients were deemed "uncensored" if they died of breast cancer within 30 years of their initial date of diagnosis.

Results and Discussion

Validation of Microarray Cohort. To validate our tissue microarray cohort of 300 node-positive breast cancers, we assessed several traditional histopathological markers of malignancy. Using univariate analysis of long-term disease-related survival, we found that large tumor size, high nuclear grade, low estrogen receptor expression, and high number of involved lymph nodes were all significant predictors of poor outcome (Table 1). We next assessed the prognostic power of HER2 immunohistochemistry, using standard brown staining, visual examination by a pathologist, and scoring on a four-point scale (0 to 3+). Manual analysis showed a typical pattern of HER2 expression with 15% of tumors overexpressing the antigen (2+ and 3+; Fig. 1B). As expected, high-level (3+) tumors showed a significantly worse outcome with a relative risk of 2.25 ($P = 0.0007$; Table 1). Analysis of HER2 gene amplification by FISH was not predictive in our study, but this was most likely attributable to the relatively small number of cases that, for technical reasons, were scorable (125 of 300; Table 1). However, both automated and manual analyses of HER2 protein levels were highly correlated with HER2 gene amplification ($P < 0.0001$). The percentage of HER2-amplified cases in each manual category were 4.0% (0), 13.7% (1+), 71.4% (2+), and 75.0% (3+) and in each AQUA category were 9.5% (normal), 13.7% (intermediate), and 77.8% (high).

HER2 Expression on Normal Epithelium. We then assessed the level of HER2 expression on normal breast epithelium with use of automated analysis on a microarray. This epithelium was derived from normal ducts and/or lobules isolated from uninvolved breast tissue taken from 84 breast cancer patients. Consistent with previous studies using biochemical assays, our results demonstrated a low but detectable level of HER2 in normal epithelium, which was tightly grouped into a single peak with a mean of 5 and a SD of 1.5 (AQUA score; Fig. 1A; Ref. 6).

Automated Analysis of HER2 Expression in Breast Cancer. In contrast to the tightly grouped peak in normal epithelium, HER2 expression in breast tumors was broadly distributed (Fig. 1C). Expression levels of HER2 in tumors exhibited a mode similar to that of normal epithelium, but with significant skew toward higher-level expression. Examination of the histogram suggested that there were three naturally occurring populations based on HER2 expression: normal, intermediate, and high (Fig. 1C). A discernable break in the histogram at AQUA score 25 divided HER2-high from the remaining tumors. The remaining tumors could then be subdivided into HER2-low and HER2-intermediate groups depending on whether their expression levels were greater than the mean HER2 expression on normal epithelium + 1 SD (AQUA score < 6.5; Fig. 1, A and C). On the basis of these divisions, 17.5% of the tumors were designated HER2 normal, 71.3% were HER2 intermediate, and 11.2% were HER2 high.

Comparison of Manual and Automated Techniques. We then compared HER2 expression as gauged by automated and manual techniques (Fig. 1, panels C and B, respectively). In contrast to AQUA scores, which were continuously scored on a scale of 0–255, manual scoring of HER2 expression was performed on a nominal four-point scale (0 to 3+). Despite this difference, regression analysis demonstrated good correlation between the two methods ($r = 0.704$). However, there was a significant degree of overlap in the automated scores of cases from adjacent manually determined groups (Fig. 1D). Whereas there was a clear division between the histograms of tumors scoring 0/1+ and 2+/3+, the distinction between tumors scoring 0 and 1+ was indistinct. This result shows the difficulty in manually translating a biological (continuous) marker into a nominal four-point scale. Even for the trained eye of a pathologist, accurate distinction between nominal categories (e.g., 2+ versus 3+) is difficult and often arbitrary. Indeed, recent studies have demonstrated a significant lack of reproducibility in the clinical determination of HER2 levels attributable in part to this difficulty (7–9).

Examination of manual and automated techniques revealed that both were equally able to define a population of tumors expressing high levels of HER2 with poor outcome (relative risk, 2.25 and 2.18; $P = 0.0007$ and 0.0013, respectively; Table 1). However, unlike manual analysis, automated analysis revealed that tumors expressing normal levels of HER2 also showed a significantly worse outcome (relative risk, 1.71; $P = 0.0091$; Table 1). Given the amount of overlap in the 0 and 1+ categories from manual scoring (Fig. 1D), it is not surprising that manual assessment of stained slides has not previously identified the HER2-normal population.

Defining the Subpopulation of HER2-normal Tumors. To determine whether HER2 expression correlated with known prognostic markers in our cohort, we assessed possible associations between HER2 and hormone receptor status, tumor size, and nuclear grade. High-level HER2 expression was correlated with high nuclear grade and inversely correlated with estrogen receptor status (Table 2). The HER2-normal population showed no significant correlation with nodal involvement, tumor size, or estrogen receptor, but did show an association with high nuclear grade ($P = 0.0494$; Table 2). Few of the HER2-normal tumors exhibited gene amplification (2 of 21 examined), ruling out the possibility that in tumors expressing

<table>
<thead>
<tr>
<th>Marker</th>
<th>n</th>
<th>P</th>
<th>Relative risk</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HER2 manual score</td>
<td>0.0071</td>
<td>153</td>
<td>1.00</td>
<td>0.68–1.52</td>
</tr>
<tr>
<td>1+</td>
<td>50</td>
<td>0.9383</td>
<td>1.02</td>
<td>0.68–1.52</td>
</tr>
<tr>
<td>2+</td>
<td>13</td>
<td>0.9763</td>
<td>1.01</td>
<td>0.49–5.08</td>
</tr>
<tr>
<td>3+</td>
<td>28</td>
<td>0.0007</td>
<td>2.25</td>
<td>1.41–3.58</td>
</tr>
<tr>
<td>HER2 AQUA score</td>
<td>0.0009</td>
<td>46</td>
<td>1.71</td>
<td>1.14–2.56</td>
</tr>
<tr>
<td>Normal</td>
<td>184</td>
<td>1.00</td>
<td>2.18</td>
<td>1.35–3.51</td>
</tr>
<tr>
<td>Intermediate</td>
<td>30</td>
<td>0.0013</td>
<td>1.07</td>
<td>0.60–1.90</td>
</tr>
<tr>
<td>High</td>
<td>22</td>
<td>0.8121</td>
<td>1.00</td>
<td>1.13–3.33</td>
</tr>
<tr>
<td>Nodal involvement</td>
<td>0.0279</td>
<td>68</td>
<td>1.00</td>
<td>0.75–1.55</td>
</tr>
<tr>
<td>1–3</td>
<td>54</td>
<td>0.6708</td>
<td>1.08</td>
<td>0.75–1.55</td>
</tr>
<tr>
<td>4–9</td>
<td>141</td>
<td>0.0086</td>
<td>1.62</td>
<td>1.13–3.33</td>
</tr>
<tr>
<td>≥10</td>
<td>102</td>
<td>0.0001</td>
<td>2.09</td>
<td>1.43–3.07</td>
</tr>
<tr>
<td>Tumor size (cm)</td>
<td>0.0007</td>
<td>80</td>
<td>1.00</td>
<td>0.92–1.93</td>
</tr>
<tr>
<td>&lt;2</td>
<td>53</td>
<td>0.1255</td>
<td>1.33</td>
<td>0.92–1.93</td>
</tr>
<tr>
<td>2–5</td>
<td>102</td>
<td>0.0001</td>
<td>2.09</td>
<td>1.43–3.07</td>
</tr>
<tr>
<td>Nuclear grade</td>
<td>95</td>
<td>0.0040</td>
<td>1.55</td>
<td>1.15–2.08</td>
</tr>
<tr>
<td>High</td>
<td>104</td>
<td>0.0262</td>
<td>1.41</td>
<td>1.041–1.906</td>
</tr>
</tbody>
</table>
normal levels of HER2, the HER2 gene is amplified but the HER2 protein is not detected.

**Multivariate Analysis of HER2-normal and -high Populations.** Finally, we determined whether normal or high expression of HER2 by tumors was an independent predictor of long-term disease-related survival. Combined multivariate analysis of HER2 with the traditional histopathological markers, nodal involvement, tumor size, nuclear grade, and estrogen receptor, demonstrated that both normal- and high-level HER2 expression were independently predictive of patient outcome (Table 3).

Our data suggest that HER2 divides cases of node-positive breast carcinoma into three categories: normal, intermediate, and high expressions. Tumors expressing either normal or high HER2 levels do poorly in long-term follow-up. Of particular note are three previous studies that have looked at HER2 expression levels using "gold standard" biochemical techniques (Western blots and ELISAs; Refs. 6, 10–13). Two of these studies suggested a bimodal distribution for HER2, with both low and high levels correlating with known markers of tumor aggression (10–12), but a third found no such distribution (13). Because such techniques require fresh tissue for analysis, they were unable to assess long-term follow-up on a large cohort of patients. The AQUA-based analysis provides quantitative information from tissue microarrays constructed from archival tissues; we thus were able to examine a large cohort of patients with known long-term disease-related survival. Our data show that normal HER2 expression is an independent prognostic indicator of poor outcome and demonstrate that, unlike manual immunohistochemical analysis, automated analysis can identify a patient population that is otherwise detectable only by established biochemical assays.

HER2 overexpression can induce an aggressive phenotype via the activation of downstream regulators (e.g., phosphoinositide 3-kinase, Erk/MAP kinase, and Ras; Refs. 14–16). How normal levels of HER2 could be associated with a similar aggressive phenotype is unknown at present. We speculate that these tumors might overexpress another growth factor receptor that promotes tumor aggression via a ligand-dependent or -independent mechanism. It is possible that expression

<table>
<thead>
<tr>
<th>Marker</th>
<th>All cases</th>
<th>Normal (%)</th>
<th>Intermediate (%)</th>
<th>High (%)</th>
<th>$P$ ($\chi^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td>Normal vs. intermediate</td>
</tr>
<tr>
<td>Nodes positive</td>
<td>268</td>
<td>54</td>
<td>60</td>
<td>54</td>
<td>43</td>
</tr>
<tr>
<td>1–3</td>
<td>145</td>
<td>26</td>
<td>23</td>
<td>27</td>
<td>33</td>
</tr>
<tr>
<td>4–9</td>
<td>70</td>
<td>20</td>
<td>17</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>10</td>
<td>53</td>
<td>20</td>
<td>17</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>Tumor size (cm)</td>
<td>238</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>102</td>
<td>43</td>
<td>35</td>
<td>44</td>
<td>48</td>
</tr>
<tr>
<td>2–5</td>
<td>81</td>
<td>34</td>
<td>44</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>&gt;5</td>
<td>55</td>
<td>23</td>
<td>21</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>Nuclear grade</td>
<td>269</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>97</td>
<td>36</td>
<td>45</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Estrogen receptor</td>
<td>265</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>105</td>
<td>40</td>
<td>39</td>
<td>34</td>
<td>77</td>
</tr>
</tbody>
</table>

Table 2. Distribution of prognostic markers by HER2 level based on $\chi^2$ analysis.
of such alternate growth factor receptors in some tumors results in the down-regulation of HER2 expression via a feedback mechanism, producing aggressive tumors bearing a HER2-normal phenotype. Another possible explanation for the poor prognosis of HER2-normal tumors is that high levels of coreceptor ligand-independent activation of HER2 might result in the internalization and degradation of the receptor, producing apparent low-level HER2 expression. Finally, HER2-normal breast cancers may represent a population of aggressive poorly differentiated neoplasms that have developed HER2- and growth factor-independent mechanisms for their growth. The association between normal HER2 expression levels and high nuclear grade supports this idea. Recent data from the Brown and Botstein group also support this finding. They showed five unique breast cancer classes by cDNA array clustering experiments, two of which had very poor outcomes. One of these groups was HER2 positive, but the other showed no evidence of HER2 overexpression (17).

From a clinical perspective, response to Herceptin has largely been seen in HER2 high expressers or HER2-amplified cases. This may be attributable to the fact that 2+ or 3+ levels of expression were required for entry into most clinical trials (18–20). The response of 0 or 1+ tumors to paclitaxel with and without Herceptin is being studied in a large randomized trial (CALGB 9840; Ref. 21). Although patients with HER2-normal tumors are unlikely to respond to Herceptin, they may benefit from more aggressive traditional chemotherapy. The ability to accurately distinguish between HER2-normal and HER2-intermediate tumors by automated analysis not only has prognostic value but may also help in the development and evaluation of new therapeutics targeted to treat this subpopulation.

Acknowledgments

We thank Michael DiGiovanna and D. Craig Allred for critical review of this manuscript. We also thank Nicole Parisot, Thomas D’Aquila, Mary Helic, Lori Charette, and Diana Fischer for assistance in this effort.

References