Review: Significance of, and Optimal Screening for, HER-2 Gene Amplification and Protein Overexpression in Breast Carcinoma

William K. Funkhouser and Kathleen Kaiser-Rogers
Department of Pathology and Laboratory Medicine, University of North Carolina School of Medicine, Chapel Hill, North Carolina

Abstract. Breast carcinoma is a common disease, with an estimated 183,000 new cases expected in the USA during 2000. Whereas early stage patients have high likelihood of cure, only 20-40% of patients with metastatic breast carcinoma respond to presently available chemotherapy. A need exists to identify the underlying biological subsets of morphologically similar carcinomas in order to develop customized therapies for patients who require chemotherapy. The HER-2 receptor tyrosine kinase is overexpressed in 15-30% of breast carcinomas, and is associated with a worse prognosis stage-for-stage. Humanized monoclonal antibody therapy (Herceptin™; Genentech Co.) appears to benefit this subset of patients by improving their response rate and survival following anthracycline- or taxane-based chemotherapeutic regimens. Both HER-2 gene amplification and protein overexpression correlate with clinical outcomes, and screening for HER-2 gene amplification appears to be the more informative test. This article reviews data on the HER-2 gene and protein, discusses their association with clinical outcomes, and proposes a strategy for screening for HER-2 excess in formalin-fixed specimens of breast carcinoma. (received 30 June 2001; accepted 20 August 2001)

Keywords: breast cancer, HER-2 gene amplification, HER-2 protein overexpression, cancer chemotherapy

Introduction

Adenocarcinoma of the breast (“breast carcinoma”) has shown a persistent increase in incidence during the last two decades, with an estimated 183,000 new cases expected in the USA during 2000 [1]. Five-year survival rates over this time (all stages combined) have shown small but significant improvements, and are now at 85%. In spite of incremental improvements in detection and therapy, approximately 41,000 US women died from breast carcinoma during 2000. There is room for improvement in the customization of therapy for individual patients with newly diagnosed breast carcinoma. This article seeks to clarify the prognostic and therapeutic value of assays for the HER-2/neu proto-oncogene and its p185 gene product in newly diagnosed breast carcinomas, and to identify the optimal screening method for detection of HER-2 excess.

In addition to the well-recognized independent prognostic variables of tumor size, grade, and stage [2], HER-2 gene amplification and protein overexpression have been correlated with reduced survival in untreated patients and with improved responsiveness to some chemotherapeutic agents in treated patients [3-5]. HER-2 gene amplification and protein overexpression are observed in 15-30% of infiltrating ductal breast carcinomas [6-10]. Recent data on patients with metastatic breast carcinoma treated with the humanized murine 4D5 anti-HER-2 antibody, Herceptin™ (trastuzumab), have shown that antibody blockage of the HER-2 protein is associated with improved survival in women treated with either anthracycline- or taxane-
based chemotherapeutic regimens [11,12]. These data indicate that HER-2 as a molecular biological variable can independently predict the likely natural history of breast carcinoma, and can also, like the estrogen receptor, serve as a target for therapeutic intervention.

HER-2 protein is one of 4 members belonging to the human epidermal growth factor receptor (EGFR) family of transmembrane receptor tyrosine kinases, or type I growth factor receptor family [13]. The prototype member of this family, EGFR, was discovered first in mice [14] and later in humans [15]. EGFR was identified as a plasma membrane protein with tyrosine kinase activity [16,17]. The gene for EGFR was cloned and initially found to be amplified in epidermoid (squamoid) carcinoma cell lines [18]. Protein sequence similarity with a viral oncogene, v-erbB, was noted [19], and an erbB probe was used to identify a second human gene called c-erbB-2 or HER-2 [20,21].

The homologous protein and gene in rat, neu, had been previously identified in a mutagen neural tumor model [22, 23], hence the nomenclature for this second member, c-erbB-2/HER-2/neu. Two additional members of the EGF gene family were subsequently identified [24,25]; to date 4 members have been described for humans. They are presently referred to as Human Epidermal growth factor-like Receptors, or HER 1-4.

The human HER-2/neu gene is located on 17q21 [21], and encodes a 185kD glycoprotein which contains an extracellular domain, a transmembrane domain, and an intracellular tyrosine kinase domain [26]. The HER-2 glycoprotein forms heterodimers in vitro with the other EGFR family members [27-30]. HER-2 protein heterodimerization in in vitro model systems contributes to increased EGF and neuregulin ligand binding, phosphorylation, and increased signalling potency [4,13,31,32].

Overexpression of wildtype human HER-2 protein appears sufficient for NIH 3T3 transformation [33,34]. Anti-HER-2 monoclonal antibody therapy of HER-2-transformed NIH 3T3 allografts in nu/nu mice inhibited tumor growth [35], arguing that HER-2 activity plays a critical role in transformation and maintenance of the neoplastic proliferation. However, it is unclear at this time whether overexpression of the wildtype human protein HER-2 is sufficient for human breast epithelial transformation, or whether non-synonymous point mutations in the transmembrane domain [36] or short in-frame deletions in the extracellular domain [37] may contribute to constitutive HER-2 phosphorylation and activation. Few or no cases without HER-2 gene amplification show increased HER-2 mRNA abundance [38].

At present, overexpression of the protein and subsequent cell transformation are attributed to amplification of the wildtype gene. Given the prognostic and therapeutic significance of HER-2 excess, we sought to clarify optimal screening methods for newly diagnosed breast carcinomas.

Possible HER-2 screening methods

The goal of screening new cases of breast cancer is to identify the patient subpopulations of interest, whether as a poor prognostic subset, or as a candidate subset for customized therapies such as anti-receptor antibody therapy or receptor tyrosine kinase inhibitor therapy. HER-2 gene amplification and HER-2 protein overexpression have both been associated with poor prognosis in women with breast carcinoma [5,39]. Furthermore, patients whose metastatic carcinomas show significant HER-2 gene amplification and protein overexpression benefit from the addition of anti-HER-2 antibody therapy to conventional chemotherapy regimens [11].

The presence of covarying excesses of both the HER-2 gene and its encoded 181 kd glycoprotein raises the question of whether gene amplification or protein over-expression should be assayed as the gold standard for considerations of prognosis and potential response to antibody therapy in patients with newly diagnosed breast carcinoma.
**The HER-2 gene**

Fifteen to 30% of breast carcinomas are found to have associated amplification of the HER-2 gene on chromosome 17q21 [6-9]. Aneuploidy of chromosome 17 is observed in 4-25% of invasive breast carcinomas [10,40,41]; its significance with regard to prognosis and treatment is unknown. Genbank and Medline searches reveal no published records of human HER-2 gene insertions, deletions, or base substitutions, suggesting that amplified human HER-2 genes contain few if any mutations or polymorphisms. These data suggest that amplification involves the wildtype allele, and that the biological effects associated with human HER-2 gene amplification are due to excess wildtype HER-2 protein. Ideal nucleic acid screening methods would therefore require the ability to detect human HER2 gene amplification, possibly normalized to chromosome 17 ploidy level.

A priori, DNA-based methods that would meet these requirements would include Southern blots, quantitative PCR, and in situ hybridization. The optimal assay must be able to evaluate scant amounts of invasive carcinoma within formalin-fixed breast core biopsies containing a heterogenous mixture of normal and cancerous cells. The small amount of carcinoma DNA in tissue core biopsies, and its dilution by normal stroma and in situ carcinoma, makes Southern blotting an unsatisfactory assay for detection of HER2 gene amplification. Although quantitative PCR is ideal for small samples, and a short amplicon product assay could be designed to compensate for formalin-induced template DNA nicking, quantitative PCR would not overcome the template dilution effect of normal stroma and in situ carcinoma, without a time-consuming microdissection purification step such as laser capture microdissection [42].

In situ hybridization targeting the HER2 gene is the one method that could be performed on scant biopsy material, compensate for formalin-fixed template DNA nicking, and generate results that are specific to the invasive carcinoma component. Although chromogenic ISH methods analogous to immunohistochemistry are under development, fluorescence in situ hybridization (FISH) is now available. This method allows the number of HER2 genes to be scored specifically within the invasive carcinoma cells, and generates accurate tumor population estimates from as few as 20-60 cells [43].

Since 4-25% of breast carcinomas show aneuploidy for chromosome 17 [40,41], normalization of HER-2 copy number to chromosome 17 copy number could be needed to correct for hyperdiploidy of chromosome 17. Two-probe, two-color measurements of the ratio of HER-2 copies to an internal chromosome copy number control (eg, the chromosome 17 centromere) have been developed [44] and are commercially available (PathVysion, Vysis Co.). This two-color method shows good interobserver reproducibility of scoring and allows normalization for chromosome 17 ploidy level, both within a heterogeneous carcinoma and between patients. Although normalization for chromosome 17 ploidy has been shown in some studies to correlate better with overall and disease-free survivals than raw HER-2 copy number [39], a need for this internal control copy number standard is not universally accepted [38]. There is, however, general agreement that the chromosome 17 centromeric probe provides an internal control for hybridization efficiency [10,38]. Therefore, at present, the optimal screening method for HER-2 gene amplification in small formalin-fixed biopsies appears to be the chromosome 17-normalized FISH assay.

The HER-2 gene encodes a full-length 4.6 kb transcript. A 2.3 kb transcript that would translate into only the extracellular domain (ECD) has been observed in some human breast carcinoma cell lines, presumably as the result of alternative splicing [45].HER-2 mRNA abundance correlates with underlying HER-2 gene amplification [7,38], so that Northern blotting or quantitative RT-PCR methods could identify and quantify gene amplification.
However, the same potential sources of inaccuracy discussed above for DNA blotting and PCR assays on a heterogeneous population apply. In addition, RNA is less stable than DNA in fresh and fixed tissue, so that mRNA levels in formalin-fixed, paraffin-embedded tissue may bear no correlation with in vivo mRNA levels. The strengths and weaknesses of the FISH, PCR, Southern blots, and Northern blots, as described above, support the conclusion that the chromosome 17-normalized FISH assay is presently the optimal method for screening for HER-2 gene amplification in small formalin-fixed biopsies.

**The HER-2 protein**

HER-2/neu is a 185 kD glycoprotein [22] that is overexpressed in from 15-40% of breast carcinoma cases [10,46]. HER-2 protein can be detected by Western blotting or immunohistochemistry. Immunohistochemical detection of HER-2 protein in breast carcinoma shows marked variation among antibodies [47]. As discussed above for the nucleic acid-based methods, the results obtained from a Western blot represent a mixed population of normal, in situ carcinoma, and invasive carcinoma cells, rather than just invasive carcinoma cells. The presence of normal cells within the tested population would tend to normalize or dilute the result obtained by this method.

In contrast, immunohistochemistry has the specific ability to detect the HER-2 protein and assign its expression to the invasive carcinoma component. Numerous monoclonal and polyclonal antibodies capable of recognizing HER-2 in paraffin sections by immunohistochemistry have been reported in the literature. Unfortunately, the sensitivity of these antibodies appears to be markedly variable, so the published data regarding correlations between HER-2 protein expression, HER2 gene amplification as estimated by FISH, and overall or disease-free survival are confusing and sometimes difficult to interpret [39, 47-49]. Similarly, poor correlation between data for HER-2 immunoreactivity and estimates of FISH gene amplification has also been attributed to differences in antibody [47]. Differences in the methods for fixation [50] and antigen retrieval [51], absence of gold standard diagnostic criteria for immunostained sections, and poor interobserver reproducibility of interpretations have also been cited [5,9,10,32,52].

Such criticisms can be addressed by generating a set of agreed-upon antigen retrieval, staining, and scoring criteria for a single antibody, as has been done for the Herceptin™ clinical trials monoclonal antibodies 4D5 and CB11 [53-56] and for the FDA-approved polyclonal goat anti-hHER-2 antibody A4085 (HercepTest™, DAKO Co.). In recent studies, the HercepTest™ has yielded good interobserver scoring reproducibility and high sensitivity for HER-2 in paraffin section immunohistochemistry [52]. However, compared to other studies using various commercial reagents, the HercepTest™ appears to identify a higher percentage of patients with HER-2 overexpression (2+ or 3+ by DAKO criteria; 35-40% with HercepTest™ versus 19-27% with other tests) [38, 42,49,57]. These data, in addition to the poor correlation between HercepTest™ 2+ reactivity and HER-2 mRNA abundance, noted by Tubbs et al [38], suggest that many HercepTest™ 2+ (“weakly overexpressed”) cases are being overinterpreted as showing overexpression (ie, false positive results).

**Correlation of HER-2 gene amplification and protein overexpression**

There appears to be a good, but not perfect, correlation between normalized FISH estimates of gene amplification and paraffin immunohistochemistry estimates of protein overexpression [38, 39,58]. Slamon et al [7] showed 90% concordance in amplified breast carcinomas between gene amplification and gene expression using Northern blots, Western blots, and immunohistochemistry.

Hoang et al [9] showed a 90% concordance between HercepTest™ and FISH, with high (0.78
kappa statistic) interobserver HercepTest™ scoring reproducibility. In this series, 16 of 17 cases with HER-2 gene amplification showed 3+ HercepTest™ reactivity, while 5 of 7 cases without HER-2 gene amplification showed 2+ HercepTest™ reactivity. These data suggest that the correlation between gene amplification (as assayed by FISH) and protein overexpression (as measured by immunohistochemistry) is strong when the HercepTest™ score is high (3+), and weak when the score is only 2+.

Similarly, Jacobs et al [59] found that 75% of their discordant cases (immunohistochemistry positive and FISH negative) had a HercepTest™ reactivity score of 2+. Tubbs et al [38] published additional data that argue that a HercepTest™ score of 2+ may not be a satisfactory predictor of underlying HER-2 gene amplification and should not be used as a criterion for Herceptin™ candidacy. When their results were stratified by HercepTest™ %o score, only 17% of the CB11 and HercepTest™ 2+ cases showed HER-2 gene amplification, and all of their HercepTest™ 2+ “false positive” (FISH negative) cases were also mRNA negative by RNA:RNA in situ hybridization. These data suggest that few if any cases of HER-2 protein overexpression are due to transcriptional upregulation of HER-2 without HER-2 gene amplification, and that when discordant HER-2 FISH and HercepTest™ results are obtained, the FISH results may be more biologically accurate.

Series using quantitative PCR to detect HER-2 gene amplification showed weak CB11 immunohistochemical reactivity in 6% of unamplified cases [60], and “weak-strong” TAB250, CB11, and HercepTest™ reactivity in 5% of unamplified cases [42]. In the latter series, the HercepTest™ showed higher sensitivity, lower specificity, and lower positive predictive value than the other two antibodies [42]. In addition to these series showing discordant overexpression of the HER-2 protein, several series contain some cases with discordant underexpression of HER-2 protein as estimated by the R60, CB11, TAB250, or HercepTest™ %o reagents [38,39,42].

Clearly, there is significant variation in the correlation between HER-2 gene amplification and the different commercial antibody estimates of HER-2 protein overexpression, even when using the Dako HercepTest™ with standardized antigen retrieval, staining, and grading criteria. It is unclear whether these results indicate false results for protein expression or for gene amplification. The decision whether protein overexpression or gene amplification should serve as a gold standard for considerations of patient prognosis, chemotherapy choice, and antibody therapy depends on clinical outcome correlations, discussed below.

HER-2 as a prognostic variable

Given that both HER-2 gene amplification and protein overexpression screening methods are technically and economically feasible, we question which is the more informative method for screening newly diagnosed breast carcinomas. The following discussion reviews the published data on how these two methods correlate with clinical outcomes, both as a prognostic variable of natural history and as a predictive variable of therapeutic response.

Two publications address the issue of whether gene amplification or protein overexpression should be used as the gold standard for case prognosis and anti-HER-2 antibody therapy consideration. Both studies make an argument that gene amplification is the more robust prognostic variable. The first publication, by Ross and Fletcher [5], inventories the pre-Herceptin™ study data on HER-2 and clinical outcomes. Of 52 studies describing almost 17,000 patients with invasive breast carcinoma, 33 studies (63%) describing 11,900 patients (70%) showed either gene amplification or protein overexpression to be a significant independent prognostic variable by multivariate statistical analysis. Those studies that failed to show multivariate independent prognostic significance were more likely to have been based on immunohistochemistry than on gene-based screening methods.
The second study, by Pauletti et al [39], used paired HER-2 FISH and immunohistochemical data from 856 patients to determine which serial test order was most informative regarding clinical survival. They showed that immunohistochemistry results of 2+ or 3+ gained additional prognostic significance when followed by FISH analysis. The converse was not true; no additional prognostic significance was gained when FISH (+) cases were subsequently studied by immunohistochemistry. This study was performed using the R60 polyclonal antibody, but good correlation between their R60 and DAKO HercepTest™ results was observed within a small sample set.

A possible explanation for the apparent diluted prognostic value of HER-2 immunohistochemistry was suggested by Birner et al [61]; in their study, HercepTest™ 3+ reactivity correlated with survival, whereas HercepTest™ 2+ reactivity did not, suggesting that HercepTest™ 2+ cases should not be considered abnormally overexpressed.

Although still inconclusive, the data described above strongly support the argument that gene amplification screening methods may be more prognostically accurate than immunohistochemical screening methods.

**HER-2 as a predictor of response to therapy**

Studies have suggested that HER-2 immunophenotype correlates with relative resistance to some chemotherapeutic agents (eg, Methotrexate or Tamoxifen) [62,63], and to relative sensitivity to others (eg, Doxorubicin) [64,65]. Phase II trials with the anti-HER-2 humanized 4D5 antibody Herceptin™ used 4D5 or CB11 monoclonal antibody reactivity with two different sets of scoring criteria to define trial candidacy, and found a 12-25% overall response rate [53-56]. The most prominent response rate was found among the strongly HER-2 immunoreactive (3+) cases [56,66].

Data on 459 patients with late stage HER-2-overexpressing breast carcinoma demonstrated that patients who received the anti-HER-2 humanized antibody Herceptin™ in addition to either of two conventional chemotherapy regimens had a longer duration of response, an extended period prior to progression, a higher 1-yr survival rate, and longer overall survival[11]. Side effects noted in the study included cardiac dysfunction in 27% of patients receiving both the antibody and anthracycline, which reinforces the need for accurate assessment of patient candidacy for anti-HER-2 antibody therapy.

The best screening method for HER-2 excess for use in predicting Herceptin™ response remains to be addressed in the literature. Nonetheless, the data that have been discussed demonstrate that HER-2 protein overexpression can serve as a predictive variable for response to therapy, as well as a target for customized antibody therapy, in patients with late stage breast carcinoma.

**Conclusions**

The correlation between HER-2 excess and both a poor prognosis and a differential response to therapeutic agents suggests that routine HER-2 screening should be performed as part of the evaluation of each new case of breast carcinoma. It is not yet known whether FISH or immunohistochemistry best predicts Herceptin™ response, but the collected observations of Ross et al [5] and Pauletti et al [39] argue strongly for use of gene amplification as a gold standard for clinical outcome prediction. A good argument can therefore be made for use of two-color FISH as the sole screening method for detection of HER-2 excess.

For those labs that have chosen instead to use immunohistochemistry to screen for HER-2 excess, it appears that a high level of HER2 protein expression (eg, 3+ HercepTest™ immunoreactivity) accurately predicts HER2 gene amplification, and that scores of 0-1+ HercepTest™ immunoreactivity accurately exclude HER2 gene amplification. In contrast, 2+ HercepTest™ immunoreactivity score,
or “weak overexpression” by DAKO criteria, does not appear reliable for predicting HER-2 gene amplification, prognosis, or therapeutic response.

In view of the lack of correlation between a HercepTest™ score of 2+ and HER-2 gene amplification, and the recognition that gene amplification may be a gold standard for clinical outcome, we recommend that HER-2 gene amplification be confirmed by FISH analysis in all cases with a HercepTest™ score of 2+.

References

41. Press MF, Bernstein L, Thomas PA, et al. HER-2/


