ERBB2 and TOP2A in Breast Cancer: A Comprehensive Analysis of Gene Amplification, RNA Levels, and Protein Expression and Their Influence on Prognosis and Prediction

Jan C. Brase¹, Marcus Schmidt², Thomas Fischbach², Holger Sültmann¹, Hans Boja³, Heinz Koelbl², Birte Hellwig³, Jörg Rahnenführer³, Jan G. Hengstler³, and Mathias C. Gehrmann³

Abstract

Purpose: The prognostic and predictive relevance of epidermal growth factor receptor 2 (ERBB2) and topoisomerase II α (TOP2A) have long been a matter of debate. However, the correlation of DNA amplification, RNA levels, and protein expression and their prognostic role and association with anthracycline responses in node-negative breast cancer have not yet been evaluated.

Experimental Design: We first analyzed TOP2A and ERBB2 at the levels of gene amplification, and RNA and protein expression, and studied their correlations. Additionally, TOP2A and ERBB2 were analyzed in 782 node-negative breast carcinomas in patients who did not receive systemic therapy and in 80 patients treated with epirubicin and cyclophosphamide (EC) prior to surgery.

Results: TOP2A gene amplification did not correlate with protein expression (P = 0.283) and showed an association with gene expression with only borderline significance (P = 0.047). By contrast, TOP2A RNA levels correlated with protein expression (P < 0.001). TOP2A gene expression was significantly associated with the metastasis-free interval (MFI; P < 0.001) and was associated with complete remission in patients treated with EC (P = 0.002). In contrast to TOP2A, ERBB2 gene amplification correlated with RNA level (P < 0.001) and protein expression (P < 0.001). ERBB2 gene expression was associated with the MFI only in estrogen receptor–positive carcinomas, whereas ERBB2 protein expression (P = 0.032) was associated with MFI in the entire cohort.

Conclusions: Overall, our study indicates that the TOP2A RNA level is a good prognostic marker and is also associated with a favorable response to anthracyclin-based therapy. By contrast, ESR1 was associated with poorer responses to anthracyclin-based therapy, whereas the association with ERBB2 RNA was not significant. Clin Cancer Res; 16(8); 2391–401. ©2010 AACR.
relationships between gene amplification, RNA levels, and protein expression of both factors are not fully understood. Therefore, we conducted a pilot study to determine DNA amplification, and gene and protein expression of both ERBB2 and TOP2A in 53 primary breast carcinomas. Next, we studied their prognostic relevance at the RNA level in 782 node-negative breast carcinomas from patients who did not receive systemic adjuvant therapy. Finally, we analyzed TOP2A and ERBB2 in a third cohort of 80 patients who received neoadjuvant chemotherapy with epirubicin and cyclophosphamide. TOP2A RNA levels were significantly associated with the metastasis-free survival intervals in node-negative tumors. In contrast to ERBB2, TOP2A was associated with responses to anthracyclin-based therapy. Our study suggests that TOP2A gene expression is adequate to predict metastasis-free survival and that it is associated with responses to anthracyclin-based therapy.

Materials and Methods

Pilot study cohort (n = 53). The pilot study for analysis of correlations among gene amplification, RNA level, and protein expression was done on breast cancer tissue specimens of a previously described cohort of consecutive patients who underwent surgery for primary breast cancer at the Department of Obstetrics and Gynecology, University of Mainz, Medical School, between October 2002 and September 2003 (14). The clinicopathologic characteristics are summarized in Supplementary Table S1A.

FISH, quantitative reverse transcriptase-PCR, and immunostaining used in the pilot study cohort. FISH analysis for TOP2A and ERBB2 amplification was done as described by Schmidt et al. (15). Briefly, a dual-color probe (Dakocytomation) containing a spectrum orange-labeled TOP2A gene (17q21-q22) or ERBB2 gene (17q11.2-q12) probe and a spectrum green-labeled centromere control for chromosome 17 (17p11.1-q11.1) were employed. Appropriate positive controls were included in each staining run. Histologic analysis was done with a Zeiss fluorescence microscope (Axioskop 2), whereby the samples were blinded with respect to the clinical outcome. A minimum of 80 nonoverlapping nuclei were evaluated and the ratios of TOP2A or ERBB2 signals per nucleus relative to chromosome 17 centromere signals were calculated. Ratios of ≥2 were classified as TOP2A or ERBB2 amplification. Tumors with an ERBB2 immunohistochemistry score of 2+ and ERBB2 amplification were finally considered ERBB2 positive.

Total RNA was isolated from 5-μm sections of formalin-fixed, paraffin-embedded tumor tissues after histopathologic confirmation of a tumor cell content of ≥70%. Reverse transcriptase-PCR (RT-PCR) was done with primers and probes for ERBB2 and TOP2A designed by using the Primer Express software (Applied Biosystems).

The sequences of probe, forward and reverse primer for (a) TOP2A and (b) ERBB2 were:

(a)
Probe: CAGATCAGGAACAGATGTTCCACAT
Forward: CATTGAAGCCTGCTGATATGG
Reverse: CCGTTGTGATGGGAAAAATTAATCG

(b)
Probe: ACCAGACCCACAGAGCCGG
Forward: CCAGCTTCGCCAACCTCTAT
Reverse: TGCCCTAGGTGTCCTTTT

To standardize the amount of sample RNA, GAPDH was selected as a reference gene and relative gene expression was calculated. Primers and probes were obtained from Eurogentec. Immunohistochemical analyses were done as described by Schmidt et al. (15). Briefly, 4-μm-thick paraffin sections were prepared according to standard procedures. Serial sections of formalin-fixed and paraffin-embedded tumor tissues were stained with monoclonal TOP2A antibodies (M7186, Dako Glostrup) and polyclonal ERBB2 antibodies (A0485, Dako). The sections were incubated with a biotin-labeled secondary antibody and streptavidin-peroxidase for 20 min each, then for 5 min with 0.05% 3’-diaminobenzidine-tetrahydrochloride and, finally, lightly counterstained with hematoxylin (16). All series included appropriate positive and negative controls. The immunohistochemical evaluation was done whereby the samples were blinded with respect to the clinical outcome. ERBB2 was scored from 0 to 3+ (17, 18). TOP2A was evaluated using an immunoreactive score defined by the product of a proportion score (0, none; 1, <10%; 2, 10-50%; 3, 51-80%; 4, >80% positive cells) and an intensity score (0, no staining; 1, weak; 2, moderate; 3, strong). Furthermore, a novel TOP2A status similar to the established ERBB2 scoring system was used to classify TOP2A overexpression. Only tumors with strong specific TOP2A staining (intensity score 3) in >10% of the tumor cells were considered as TOP2A status positive.
Study cohort for analysis of prognosis \((n = 782)\). The Mainz study cohort \((19)\) consisted of consecutive lymph node-negative breast cancer patients treated at the Department of Obstetrics and Gynecology of the Johannes Gutenberg University Mainz between 1988 and 1998 for whom fresh frozen tissue was available \((n = 194)\). Patients of the Mainz cohort were all treated with surgery and did not receive systemic therapy in the adjuvant setting. The established prognostic factors (histologic grade, tumor size, age at diagnosis, and steroid receptor status) were collected from the original pathology reports of the gynecologic pathology division of the University of Mainz \((20)\). All tumor samples were snap-frozen and stored at \(-80^\circ\)C. Gene expression profiling of the patients’ RNA was done using the Affymetrix HG-U133A array and the GeneChip System, as described previously \((19)\). ERBB2 and TOP2A expression levels were determined using Affymetrix probe set 216836_s_at and 201292_at/201291_s_at, respectively.

Results obtained from the Mainz cohort were validated in two previously published microarray data sets. Two further breast cancer Affymetrix HG-U133A microarray data sets were downloaded from the National Center for Biotechnology Information GEO data repository \((n = 588)\). The first data set, the Rotterdam cohort \((21)\), represents 180 lymph node-negative relapse-free patients \((GSE2034)\) and 106 lymph node-negative patients who developed distant metastases. None of these patients had received systemic neoadjuvant or adjuvant therapy. The second data set, the Transbig cohort, consisted of 302 samples from breast cancer patients that remained untreated after surgery \((22, 23)\). GEO Sample (GSM) accession numbers of samples \((GSE6532 \text{ and } GSE7390)\) used for analysis are listed in the supplementary tables published by Schmidt et al. \((19)\). In all cases, raw.cel file data were processed by MAS 5.0 using a target intensity \((TGT\) value\) of 500. The clinicopathologic characteristics of the Mainz, Transbig and Rotterdam subcohorts have been summarized in Supplementary Table S1B to D. The median follow-up periods for the Mainz, Transbig, and Rotterdam cohorts were, 7.78, 9.58, and 7.17 y, respectively.

Neoadjuvant study cohort \((n = 80)\). To evaluate the association between ERBB2 and TOP2A and the response to anthracycline therapy, we studied a cohort of 80 patients (clinicopathologic characteristics in Supplementary Table S1E) who received neoadjuvant chemotherapy with EC. EC consisted of epirubicin 90 mg/m\(^2\)/d in a short i.v. infusion, and cyclophosphamide 600 mg/m\(^2\)/d in a short i.v. infusion. Four cycles of EC were administered. Some patients received additional tamoxifen, letrozol, or seldom goserelin, for 4 to 5 wk after the EC course and before surgery. All tumor samples were collected as needle biopsies of primary tumors prior to any treatment. The biopsies were obtained under local anesthesia using Bard MAGNUM Biopsy Instrument \((C.\ R.\ Bard,\ Inc.)\) with Bard Magnum biopsy needles \((BIP\ GmbH)\), using ultrasound guidance. Samples were collected according to routine procedures for pathologic diagnosis and following institutional review board guidelines. Pathologic examination was carried out for all tumor samples by the same pathologist at the Interdisciplinary Breast Center IBC. The remainders of the samples were snap-frozen.

Total RNA was extracted from cell lysates of ground tissue, and gene expression profiling of the patients’ RNA was done using the Affymetrix HG-U133A array and the GeneChip System, as described previously \((24)\). After washing and staining, arrays were scanned using a Gene Array scanner 2500 \((Affymetrix)\). ERBB2 and TOP2A expression levels were determined using Affymetrix probe set 216836_s_at and 201292_at/201291_s_at, respectively.

Pathologic complete remission was defined as an absence of invasive carcinoma in the breast (assessed by the examining pathologist) and a lack of lymph nodal involvement. Clinical complete remission was defined as a clinical absence of invasive carcinoma of the breast. This parameter was used as a surrogate for pathologic complete remission on one occasion when a patient declined post-neoadjuvant surgical excision. Partial responder was defined as a reduction in the tumor mass of both perpendicular dimensions ranging from 10% to 75% of the initially measured tumor size, based on dynamic contrast-enhanced magnetic resonance imaging or magnetic resonance tomography, and, in some cases, on both magnetic resonance tomography and ultrasound. Nonresponder or no change was defined as an absence of tumor reduction or an increase in tumor size (stable or progressive disease).

Statistical analysis. Survival rates were calculated according to the Kaplan-Meier method. The metastasis-free interval \((MFI)\) was calculated from the date of breast cancer diagnosis to the detection date of distant metastasis. Disease-free survival was calculated from the date of breast cancer diagnosis to the detection of local recurrence/distant metastasis/cancer of the contra lateral breast, or death from cancer. Patients who died of an unrelated cause were censored at the date of death. Survival times were compared using Kaplan-Meier analysis and the Log-rank test. Univariate and multivariate Cox survival analyses were done without variable selection. For Cox analysis TOP2A was used as a continuous variable, because it showed a unimodal distribution \((Supplementary\ Fig.\ S1)\). By contrast, ERBB2 was dichotomized using 12.6 \((arbitrary\ units)\) as a cutoff point because this value separates the two subpopulations with low versus high ERBB2 gene expression \((Supplementary\ Fig.\ S1)\). Hormone receptor status was dichotomized into positive \((estrogen\ receptor\ and/or\ progesterone\ receptor\ positive)\) and negative \((both,\ estrogen\ receptor\ and\ progesterone\ receptor\ negative)\); histologic grade in G3 versus G1 and G2; and pT stage in pT2 and pT3 \((\geq 2\ cm)\) versus pT1 \((\leq 2\ cm)\). For Kaplan-Meier plots, TOP2A was dichotomized at the median, whereas ERBB2 was dichotomized at 12.6 \((arbitrary\ units)\).

Principal component analysis \((PCA)\) was done using GeneSpring 7.0 \((Agilent\ Technologies)\). Clinical information was visualized as categorical or continuous variables,
and relative gene expression was visualized on a relative scale from red, indicating high expression, to green, indicating low expression.

For multivariate analysis of a possible association with complete response versus partial response or no change, a binary regression model was used (neoadjuvant cohort treated with EC). Correlations were calculated using the Spearman test. All P values are two-sided with a significance level of α = 0.05. The Mann-Whitney U-test was applied to compare two groups of unrelated samples. The Kruskal-Wallis test was used to compare differences between more than two groups of unrelated samples. As no correction for multiple testing was done they are descriptive measures. All analyses were done using SPSS17.0.

Results

**TOP2A RNA levels but not TOP2A gene amplification correlate with protein expression.** To analyze the correlation between gene amplification and protein expression of ERBB2 and TOP2A, we used breast cancer tissue specimens from 53 patients who underwent surgery for primary breast cancer. ERBB2 amplification was observed in 12 of the 53 tumors (22.6%). Of these, six samples (11.3%) showed an increased gene copy number of TOP2A. Only one tumor exhibited TOP2A amplification without concomitant ERBB2 amplification. Neither ERBB2 nor TOP2A amplifications were observed in the remaining 40 samples (75.5%).

Of the breast cancer specimens 22% were identified as ERBB2 positive (ERBB2 immunohistochemistry score 3+), whereas 78% were identified as ERBB2 negative (<3+). ERBB2 gene amplification showed a significant correlation with protein expression (P = 0.001). All patients without ERBB2 amplifications were negative for protein expression and 91.7% with ERBB2 amplification showed a positive ERBB2 (3+) immunostaining, suggesting a strong association between ERBB2 gene amplification and protein expression. In contrast to ERBB2, amplification of TOP2A did not correlate with immunohistochemically determined protein expression (P = 0.283). To understand why ERBB2 but not TOP2A amplification is associated with higher protein expression, we systematically studied possible correlations between DNA amplification and RNA levels, as well as between gene and protein expression. Tumors with amplified ERBB2 showed significantly higher levels of ERBB2 gene expression (P < 0.001; Fig. 1A). By contrast, TOP2A gene amplification only showed a borderline significant association with TOP2A RNA level (P = 0.047) and the difference between TOP2A amplified and nonamplified tumors was relatively small (Fig. 1B). A considerable number of tumors without TOP2A amplification expressed TOP2A RNA levels, which were in the same range as samples with TOP2A amplification.

When we studied the relationship between RNA levels and protein expression, significant correlations were observed for both ERBB2 and TOP2A (Fig. 2). ERBB2 gene expression correlated with the immunohistochemically determined score ranging from 0 to 3+ (P < 0.001; Fig. 2A) and ERBB2 status (P < 0.001; Fig. 2B). Similarly, TOP2A RNA levels correlated with immunostaining intensity (P = 0.001; Fig. 2C), as well as immunohistochemically determined TOP2A status (P < 0.001; Fig. 2D).

**Prognostic relevance of TOP2A and ERBB2.** The association between TOP2A/ERBB2 gene expression and patient prognosis was analyzed respectively using a previously described cohort of 782 node-negative breast cancer patients, consisting of three subcohorts (Mainz, Rotterdam, and Transbig). Expression levels of TOP2A were associated with the MFI in the univariate Cox model [P < 0.001; hazard ratio (HR), 1.38; 95% confidence interval (95% CI), 1.24-1.53]. Because the clinical factors – age, histologic grade, pT stage, and hormone receptor status – were only available for the Transbig (n = 302) and the Mainz (n = 194) subcohorts, multivariate Cox analysis could only be done for these patients. Nevertheless, TOP2A gene expression was significantly associated with the MFI, even after
adjusting the Cox model for age, grading, pT stage, and hormone receptor status in the combined cohort \( (n = 496; \text{Table 1A}) \), as well as in the Transbig and Mainz subcohorts (Supplementary Table S2). Kaplan-Meier analysis shows the association of high \( \text{TOP2A} \) gene expression with survival time \( (P < 0.001; \text{Fig. 3A}) \).

In contrast to the unimodal distribution of \( \text{TOP2A} \), \( \text{ERBB2} \) RNA showed a bimodal distribution: the majority of patients expressed levels <12.6 (log 2 transformed data for \( \text{ERBB2} \)) and a minority expressed higher \( \text{ERBB2} \) levels (Supplementary Fig. S1). The analysis of the entire cohort \( (n = 782) \) revealed no significant association between \( \text{ERBB2} \) gene expression and the MFI after dichotomizing the data using 12.6 as a cutoff point \( (P = 0.173; HR, 1.27; 95\% \text{ CI, 0.90-1.77}) \). In the multivariate regression model, \( \text{ERBB2} \) RNA was not associated with MFI. Additional Kaplan-Meier analysis of the total cohort of 782 patients showed that \( \text{ERBB2} \) expression >12.6 was associated only with a trend towards shorter survival times, although this was not statistically significant \( (P = 0.103; \text{Fig. 3B}) \).

To verify the findings from the combined cohort of 782 patients, the data set was divided into three individual subcohorts of 194 (Mainz cohort), 286 (Rotterdam cohort), and 302 (Transbig cohort) patients. For \( \text{TOP2A} \), significant results were also obtained in the three subcohorts (Supplementary Fig. S2). In all analyses, the follow-up period was limited to 5 years. However, similarly significant associations were obtained if a follow-up period of 10 years \( (P < 0.001) \) was applied (exemplified in the combined cohort; Supplementary Fig. S3).

**Fig. 2.** \( \text{ERBB2} \) (A, B) and \( \text{TOP2A} \) (C, D) RNA correlate with their protein expression. Significant correlations were obtained for an immunostaining score, differentiating between four distinct scores with increasing intensity (A, C), as well as the \( \text{ERBB2} \) and \( \text{TOP2A} \) status (B, D) which differentiate between positive and negative tumors. Differences in RNA levels were evaluated using the Kruskal-Wallis test (A, C) and the Mann-Whitney \( U \)-test (B, D) respectively. Gene expression was determined by quantitative RT-PCR.
Correlation of TOP2A with proliferation composite scores and prognostic influence in breast cancer subgroups. In previous studies, TOP2A was shown to be prognostic in estrogen receptor–positive but not in estrogen receptor–negative breast carcinomas (25). In addition, ERBB2-positive and triple-negative carcinomas are often considered as a separate class of breast cancer (26). To obtain a clearer view on the prognostic relationship between TOP2A and ERBB2, we applied PCA. The relative gene expression levels of TOP2A and ERBB2, as well as time to distant metastasis, were visualized in a PCA plot of 194 breast cancer samples (Mainz cohort; Supplementary Fig. S4A-C). Additionally, breast cancer molecular subgroups, defined by gene expression of ESR1 and ERBB2 (ESR1+/ERBB2- = luminal like, ERBB2+, and ESR1-/ERBB2- = basal like), were visualized in the same PCA plot to show the relationship among these subtypes, TOP2A gene expression, and patient outcome (Supplementary Fig. S4D).

In contrast to TOP2A, ERBB2 RNA was not significantly associated with outcome in the analyzed breast cancer data set. This might be expected, given the fact that triple-negative tumors have a high TOP2A, a low ERBB2 expression, and an adverse outcome (Supplementary Fig. S4).

### Table 1. Association of TOP2A and ERBB2 gene expression with breast cancer specific metastasis-free survival in patients with node-negative breast cancer

#### A. Multivariate Cox analysis of TOP2A gene expression in 496 node-negative breast cancer samples. The cohort (n = 496) is a combination of the Transbig (n = 302) and the Mainz (n = 194) subcohorts

<table>
<thead>
<tr>
<th>Prognostic factors</th>
<th>P</th>
<th>HR (95% CI)</th>
<th></th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (&lt;50 vs ≥50 y)</td>
<td>0.432</td>
<td>1.166 (0.795-1.710)</td>
<td>0.432</td>
<td>1.166 (0.795-1.710)</td>
</tr>
<tr>
<td>pT stage (≤2 cm vs &gt;2 cm)</td>
<td>0.472</td>
<td>1.113 (0.832-1.489)</td>
<td>0.472</td>
<td>1.113 (0.832-1.489)</td>
</tr>
<tr>
<td>Histologic grade (Grade 3 vs grade 1 and 2)</td>
<td>0.066</td>
<td>1.503 (0.973-2.320)</td>
<td>0.066</td>
<td>1.503 (0.973-2.320)</td>
</tr>
<tr>
<td>Hormone receptor status*</td>
<td>0.570</td>
<td>1.144 (0.719-1.820)</td>
<td>0.570</td>
<td>1.144 (0.719-1.820)</td>
</tr>
<tr>
<td>(ER or PR; negative vs. positive)</td>
<td>0.567</td>
<td>1.183 (0.711-1.967)</td>
<td>0.567</td>
<td>1.183 (0.711-1.967)</td>
</tr>
<tr>
<td>ERBB2 status (positive vs negative)</td>
<td>&lt;0.001</td>
<td>1.356 (1.149-1.599)</td>
<td>&lt;0.001</td>
<td>1.356 (1.149-1.599)</td>
</tr>
<tr>
<td>TOP2A RNA level</td>
<td>&lt;0.001</td>
<td>1.356 (1.149-1.599)</td>
<td>&lt;0.001</td>
<td>1.356 (1.149-1.599)</td>
</tr>
</tbody>
</table>

#### B. Association of TOP2A gene expression with metastasis-free survival in 782 patients with node-negative breast cancer. Analysis was done separately in ER-positive versus ER-negative, ERBB2-positive versus ERBB2-negative, and in triple-negative versus non-triple-negative tumors (triple-negative “positive” means that ER, PR, and ERBB2 are negative, whereas triple-negative “negative” comprises all other patients)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>P (n)</td>
<td>HR (95% CI)</td>
<td>P (n)</td>
</tr>
<tr>
<td>ER</td>
<td>&lt;0.001 (583)</td>
<td>1.591 (1.403-1.806)</td>
</tr>
<tr>
<td>ERBB2</td>
<td>0.687 (118)</td>
<td>0.951 (0.744-1.215)</td>
</tr>
<tr>
<td>Triple negative</td>
<td>0.921 (144)</td>
<td>1.014 (0.776-1.324)</td>
</tr>
</tbody>
</table>

#### C. Association of ERBB2 gene expression with metastasis-free survival in 782 patients with node-negative breast cancer. Analysis was done separately in ER-positive versus ER-negative, Top2A low versus high (using the median as a cut point), and in triple-negative versus non-triple-negative tumors (triple negative “positive” means that ER, PR, and ERBB2 are negative, triple negative “negative” comprises all other patients)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>P (n)</td>
<td>HR (95% CI)</td>
<td>P (n)</td>
</tr>
<tr>
<td>ER</td>
<td>0.014 (583)</td>
<td>1.708 (1.115-2.616)</td>
</tr>
<tr>
<td>Top2A (high vs low)</td>
<td>0.451 (391)</td>
<td>0.852 (0.561-1.293)</td>
</tr>
<tr>
<td>Triple negative</td>
<td>0.515 (144)</td>
<td>0.048 (0.000-446.7)</td>
</tr>
</tbody>
</table>

### Abbreviations:
ER, estrogen receptor; PR, progesterone receptor.
*The hormone receptor status is positive as soon as one of both, the estrogen (ER) or the progesterone receptor status (PR), is positive.
expression of TOP2A (Supplementary Fig. S4A, D) and other proliferation-associated genes and, as a group, have a similar poor prognosis as triple-negative samples. The group-wise analysis showed that TOP2A was significantly associated with the MFI in estrogen receptor–positive and ERBB2-negative carcinomas and as soon as the triple-negative carcinomas were excluded (Table 1B). ERBB2 was significantly associated with the MFI in estrogen receptor–positive and as soon as triple-negative carcinomas (high TOP2A RNA levels) were excluded (Table 1C).

Composite scores of a variable number of proliferation-associated genes have been shown to have a strong relevance to the outcome (27, 28). Therefore, we analyzed the correlation between TOP2A, MKI67 and two composite scores consisting of AURKA, BIRC5, MYBL2, MKI67, and CCNB1 (score 1) and RACGAP1, NEK2, CENPA, RR2M, and BUB1B (score 2), and compared their prognostic performance by univariate and bivariate Cox regression analysis in the cohort of node-negative patients who did not receive systemic adjuvant therapy (Supplementary Tables S3 and S4). The analyzed parameters correlated significantly with each other (Supplementary Table S3). However, the correlation coefficient was highest between the two composite scores \( (r = 0.91) \) and lowest between the two single genes TOP2A and MKI67 \( (r = 0.65) \). Univariate Cox regression revealed that all parameters were highly significantly associated with the development of distant metastasis (Supplementary Table S4A). Significance levels were even higher, despite a smaller sample size, when the analysis was restricted to ESR1 positive and ERBB2 negative. For the remaining 2+ patients, the decision on ERBB2 status depends on the result of FISH analysis. Only if amplifications are detectable, 2+ patients are considered ERBB2 positive. Using this classification, 26 of the patients from the Mainz cohort were ERBB2 status positive and 168 were negative. The Kaplan-Meier analysis showed shorter survival times for ERBB2-positive patients compared with ERBB2-negative patients \( (P = 0.032; \text{Fig. 4A}) \). If patients were divided into 3+ versus 0 to 2+ based on the immunohistochemistry data alone, only a trend was obtained \( (P = 0.061; \text{data not shown}) \), illustrating that the additional FISH analysis of the 2+ patients leads to a slightly improved prediction. These results show that a combination of ERBB2 immunostaining and FISH is superior over analysis of ERBB2 gene expression \( (P = 0.090; \text{Fig. 4B}) \). In the multivariate analysis, ERBB2 status was not significantly associated with the MFI in the Mainz cohort for which gene array data were available \( (n = 194) \).

**Fig. 3.** Association between TOP2A and ERBB2 RNA levels and MFI in 782 patients with node-negative breast cancer. A, shorter MFIs were related to patients with TOP2A gene expression higher than the median compared with patients with relatively low expression. B, only a nonsignificant trend towards shorter MFIs was observed for patients with high ERBB2 RNA levels. A cutoff point of 12.6 (arbitrary units) was used, as indicated in Supplementary Fig. S1A.
However, in all consecutive patients (n = 402, also including patients for whom no freshly frozen material but only paraffin material was available), ERBB2 status was significant in both the univariate Cox analysis (P < 0.001; HR, 2.26; 95% CI, 1.46-3.49) and in the multivariate regression model, adjusted for age, grading, pT stage, and hormone receptor status (P = 0.001; Supplementary Table S5). Using paraffin slices from the same group of 402 consecutive breast cancer patients, we could isolate RNA of sufficient quality from 363 patients and determine ERBB2 gene expression by quantitative RT-PCR. However, ERBB2 RNA level was not associated with the MFI in the univariate COX analysis (P = 0.947; HR, 1.01; 95% CI, 0.87-1.16). Using the same RNA for quantification of TOP2A as a positive control, a significant association with the MFI was obtained (P = 0.007; HR, 1.22; 95% CI, 1.06-1.40). The results show that in this cohort, an association between ERBB2 and prognosis could be obtained by a combination of ERBB2 immunostaining and FISH but not by RNA analysis.

**TOP2A RNA level in patients with anthracycline-based chemotherapy.** The analysis of the above described cohort of 782 patients with node-negative breast cancer showed that high TOP2A gene expression is associated with a shorter MFI. This cohort is ideal to evaluate the prognostic influence of TOP2A, because the patients did not receive systemic therapy. TOP2A gene expression was also analyzed for its association with responses to anthracycline-based therapy in a cohort of 80 patients who received neoadjuvant chemotherapy with EC. Patients with complete remission (n = 12) had significantly higher levels of TOP2A RNA than patients with no change (n = 13; P = 0.002) and patients with partial responses (n = 55; P = 0.002; Fig. 5). In contrast to TOP2A, the expression of ERBB2 RNA was not associated with the response to chemotherapy (Fig. 5). When we carried out a binary logistic regression model, differentiating between complete response versus partial response and no change, only TOP2A and estrogen receptor (ESR1) RNA levels were significant, whereas ERBB2, grading, pT stage, and age were not included into the model (Supplementary Table S6). TOP2A was associated with a higher likelihood of complete response (HR, 0.2673; P = 0.031), whereas ESR1 was associated with a poor response to chemotherapy (HR, 0.0548; P = 0.028).

**Discussion**

Identification of reliable prognostic and predictive factors is still one of the largest challenges in oncology (14, 19, 20, 29–32). Furthermore, the choice of the analytical technique and the decision whether to analyze DNA amplification, RNA, or protein are equally important. Amplification of DNA can be reliably done by FISH. However, gene amplification often does not correlate to gene or protein expression because of transcriptional or translational control mechanisms. RNA can be easily and reproducibly quantified by quantitative RT-PCR. Huge public gene array data banks of breast cancer tissue are available that allow validation of results (19, 33). From a biological point of view, analysis of protein function would be ideal. However, relatively large amounts of fresh frozen tissue are required for activity assays as well as protein quantification by quantitative immunoblots (34–36) which are often not available. Therefore, immunostaining is frequently employed, because paraffin material is available...
from almost all patients. However, interpretation of immunostaining data is often limited by interlaboratory differences and difficulties in standardization (37). Therefore, the choice of the optimal technique and level of analysis should be systematically evaluated for each relevant prognostic factor because its influence might depend on the method of testing (38).

The prognostic and predictive value of measuring ERBB2 and TOP2A expression has long been a matter of debate (39–46). To our knowledge, however, a systematic study of ERBB2 and TOP2A in the same cohorts investigating the correlation of DNA amplification, gene and protein expression, and analyzing their prognostic relevance has so far not been published. In the present study, we first analyzed the correlations between gene amplification and gene levels, as well as protein expression. Next, we analyzed the expression of TOP2A and ERBB2 in a larger cohort of 782 node-negative breast cancer patients. There was no correlation between gene amplification and protein expression of TOP2A. This is in agreement with previous studies (47, 48), suggesting that control mechanisms of gene and protein expression are more relevant for levels of protein expression than gene amplification. One well-known influencing factor is that actively proliferating cells express higher RNA and protein levels of TOP2A than do nonproliferating cells (49). Additional evidence that the determination of gene amplification is not sufficient for TOP2A gene and protein expression was obtained from the ERBB2 and TOP2A RNA frequency distributions of 782 breast cancer patients: although a bimodal distribution was obtained for ERBB2 RNA this was not the case for TOP2A. Conversely, we observed a high correlation between TOP2A RNA and protein levels.

The practical implications of these results are that (a) gene amplification analysis alone is not sufficient to study a possible prognostic or predictive effect of TOP2A, and (b) TOP2A RNA analysis may be sufficient, because it correlates well with protein expression. TOP2A RNA was significantly associated with prognosis in the univariate and in the multivariate Cox analysis, adjusted for age, grading, pT stage, and hormone receptor status. Gene expression of TOP2A was not only associated with the MFI in the total cohort of 782 patients but also was significant in each individual subcohort, including the Mainz (n = 194), Rotterdam (n = 286), and Transbig (n = 302) subcohorts.

Taken together, these results show that TOP2A RNA is a powerful prognostic marker in breast cancer. TOP2A was also analyzed in molecular subgroups defined by gene expression of ESR1 and ERBB2 (ESR1+/ERBB2- = luminal like, ERBB2+, and ESR1+/ERBB2- = basal like). In line with previous studies (25) TOP2A was prognostic in estrogen receptor–positive and ERBB2-negative carcinomas but not in estrogen receptor–negative breast carcinomas. From a technical point of view, the analysis of RNA is more reproducible and easier to do than immunostaining. Therefore, there was no practical relevance to additionally analyze TOP2A by immunostaining because RNA analysis will be our method of choice to analyze the prognostic relevance of TOP2A.

There was a good correlation among gene amplification, gene expression, and immunohistochemically determined protein expression of ERBB2. Therefore, we first focused on the prognostic relevance of gene expression. In contrast to TOP2A, ERBB2 gene expression was not associated with prognosis, either in the univariate or in the multivariate Cox model in the entire cohort of 782 patients. Our group-wise comparison, however, showed that ERBB2 gene expression was associated with MFI only in estrogen receptor–positive carcinomas. We then studied the level of ERBB2 protein. For this purpose, we used exactly the same samples from 194 patients of the Mainz cohort which had already been studied for ERBB2 gene expression. Using ERBB2 immunostaining as a measure of protein expression, we observed a significant association for ERBB2 with MFI. This was confirmed by studying the total Mainz

![Fig. 5. TOP2A gene expression is associated with the response to anthracycline-based chemotherapy. In a cohort of 80 patients with neoadjuvant EC chemotherapy, patients with complete response (pCR) expressed higher levels of TOP2A RNA than patients with no change (NC; P = 0.002) or with a partial response (PR; P = 0.002). By contrast, expression of ERBB2 was not significantly associated with response to neoadjuvant EC chemotherapy.](https://clincancerres.aacrjournals.org/content/16/8/2399/F5.large.jpg)
cohort of 402 node-negative patients (also including samples from patients for whom no fresh frozen tissue was available). In the 402 patients of the Mainz cohort, immunohistochemically determined ERBB2 expression was significantly associated with poor prognosis in the univariate and in the multivariate Cox model, which was not the case for RNA analysis. The prognostic power of the immunostaining data was further improved when 2+ tumors were also analyzed by FISH, although the improvement was relatively small.

 Anthracycline-based therapies are among the most active adjuvant treatments for breast cancer. However, because of potentially serious side effects, the prediction of drug effectiveness is important. Previous studies have recommended the use of TOP2A gene amplification as a marker for sensitivity to anthracyclines (7, 8). However, according to our results it is questionable whether FISH analysis of TOP2A is adequate to predict anthracycline chemotherapy efficacy. One reason is that TOP2A amplification does not correlate with protein expression. In this aspect, our data confirm previously published studies (47, 48). A second reason is that TOP2A amplification is almost exclusively confined to ERBB2-amplified tumors, which is also in agreement with previously published data (3, 10). Therefore, FISH analysis will not allow a differentiation between predictive effects due to TOP2A and ERBB2. However, this is relevant due to the availability of targeted therapies for tumors overexpressing ERBB2. We analyzed TOP2A and its association with responses to anthracycline-based chemotherapy. We observed that TOP2A gene expression was significantly higher in patients with complete responses than in patients with no change or who only partially responded. By contrast, ESR1 expression was associated with poorer responses to anthracycline-based chemotherapy. Therefore, analysis of RNA levels of both TOP2A and ESR1 represents a good approach to identify patients who will benefit from EC chemotherapy.

In conclusion, TOP2A RNA seems to be adequate as a prognostic marker for breast cancer, based on our cohort of 782 patients who did not receive any systemic therapy. This marker is also associated with good responses to anthracycline-based chemotherapy, as shown by our cohort with neoadjuvant anthracycline-based chemotherapy with EC.

**Disclosure of Potential Conflicts of Interest**

M.C. Gehrmann, employed by Siemens Healthcare Diagnostics Products, GmbH.

**Grant Support**

Federal Ministry for Education and Research (BMBF) for financial support (NGFN project OncoProfile).

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

Received 09/11/2009; revised 01/22/2010; accepted 02/17/2010; published OnlineFirst 04/06/2010.

**References**


34. Steiner E, Pollow K, Hasencleaver D, et al. Role of urokinase-type plasminogen activator (uPA) and plasminogen activator inhibitor type 1 (PAI-1) for prognosis in endometrial cancer. Gynecol Oncol 2008;108:569–76.


Clinical Cancer Research

ERBB2 and TOP2A in Breast Cancer: A Comprehensive Analysis of Gene Amplification, RNA Levels, and Protein Expression and Their Influence on Prognosis and Prediction

Jan C. Brase, Marcus Schmidt, Thomas Fischbach, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-09-2471

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2010/04/06/1078-0432.CCR-09-2471.DC1

Cited articles
This article cites 48 articles, 23 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/16/8/2391.full.html#ref-list-1

Citing articles
This article has been cited by 8 HighWire-hosted articles. Access the articles at:
/content/16/8/2391.full.html#related-urls

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

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

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