Q uantitative real-time RT-PCR analysis of eight novel estrogen-regulated genes in breast cancer

V. Sorbello¹, L. Fusó¹, C. Sfiligoi¹, C. Scafoglio¹, R. Ponzone¹, N. Biglia¹, A. Weisz⁴, P. Sismondi¹,², M. De Bortoli °¹

¹Laboratory of Oncological Gynecology, Institute for Cancer Research and Treatment, Candiolo, Turin
²Department of Obstetrics and Gynecological Sciences, University of Turin
³Department of Oncological Sciences, Institute for Cancer Research and Treatment, Candiolo, Turin
⁴Department of General Pathology, Second University of Naples, Naples - Italy

ABSTRACT: Background: Biological markers capable of predicting the risk of recurrence and the response to treatment in breast cancer are eagerly awaited. Estrogen and progesterone receptors (ER, PgR) in tumor cells mark cancers that are more likely to respond to endocrine treatment, but up to 40% of such patients do not respond. Here, the expression of a group of estrogen-regulated genes, previously identified by microarray analysis of in vitro models, was measured in breast tumors and possible associations with other clinicopathological variables were investigated.

Methods: The expression of CD24, CD44, HAT-1, BAK-1, G1P3, TIEG, NRP-1 and RXRα was measured by quantitative real-time RT-PCR on RNA from eighteen primary breast tumors. Statistical analyses were used to identify correlations among the eight genes and the available clinicopathological data.

Results: Variable expression levels of all the genes were observed in all the samples examined. Significant associations of CD24 with tumor size, CD44 with lymph node invasion, and HAT-1 and BAK-1 with ER positivity were found. The possible combinatorial value of these genes was assessed. Unsupervised hierarchical clustering analysis demonstrated that the expression profile of these genes was able to predict ER status with an acceptable approximation.

Conclusions: Eight novel potential markers for breast cancer have been preliminarily characterized. As expected from in vitro data, their expression is able to discriminate ER- versus ER+ tumors. (Int J Biol Markers 2003; 18: 123-9)

Key words: Breast cancer, Estrogen-regulated genes, Real-time RT-PCR, Predictive factors

INTRODUCTION

Breast cancer is a composite disease, with outcomes ranging from early recurrence and death to complete cure. Responses to chemotherapy and endocrine therapy are extremely variable and sites of metastasis unpredictable. Many pathological, genetic and biochemical markers have been studied for their ability to define prognosis or predict the response to therapy. These markers include the number of invaded lymph nodes, tumor size, age, grading, histology and steroid receptors. At present, the most potent indicator of the risk of relapse is the invasion of locoregional lymph nodes. Most patients with lymph node-negative breast cancer are cured by locoregional treatment; however, about 30% relapse because traditional histomorphological and clinical factors fail to identify high-risk patients who may benefit from adjuvant chemotherapy (1). Adjuvant endocrine therapy based on estrogen ablation reduces hormone responsive breast cancer recurrence and mortality. The presence of estrogen receptor (ER) and progesterone receptor (PgR) in tumor cells has predictive value for the response to endocrine treatment. One third of ER/PgR-positive tumors fail to respond to endocrine therapy, while 10% of receptor-negative tumors do respond (2). Many other markers have been examined for their possible association with clinically relevant subgroups of breast cancer (ErbB2/neu, myc, Ki67, p53, cyclins, keratins, etc.). All show a certain degree of association with subgroups of breast cancer but none of them is really clinically useful, with the exception of ErbB2 gene amplification, which is a necessary prerequisite for Herceptin therapy.

It is a common opinion that it is necessary to combine and integrate biomolecular markers to obtain a full description of tumor biology. The cellular and molecular heterogeneity of breast tumors and the large number of genes potentially involved in controlling cell growth, death and differentiation emphasize the importance of studying multiple genetic alterations in concert (3). It has already been shown that cDNA microarrays can reclassify breast carcinomas on the basis of variations in gene expression patterns and define clinically relevant groups.
Estrogen-regulated genes analysis in breast cancer

(4, 5). Analysis of archival biopsies of tumor tissues from cases with long follow-up is the method of choice to discover profiles correlated with the risk of recurrence. In one study, breast tumors were divided into five subtypes, each characterized by high expression of different clusters of genes (6). One finding was the division of estrogen receptor-positive tumors into at least two groups with distinctive gene expression profiles and prognosis. Likewise, a signature strongly predictive of a short interval to distant metastases (“poor prognosis” signature) was identified in lymph node-negative patients (7, 8). These studies confirmed that the combinatorial analysis of multiple gene expression data can effectively identify subgroups of tumors with different clinical behaviors.

Microarray analysis was also used to monitor changes in genomewide expression of mRNAs in ZR75.1 human breast carcinoma cells in culture, after stimulation with a mitogenic dose of 17β-estradiol (E2). In this study 344 genes were found to be significantly changed in their activity after treatment with E2 (9).

In the present work eight newly identified estrogen-responsive genes were selected among this group and their expression was quantitatively assessed by real-time RT-PCR on 18 primary breast tumor samples in order to define their association with other clinicopathological characteristics. Our results suggest that a small subset of estrogen-responsive genes is sufficient to identify the ER-positive phenotype in breast cancer.

MATERIALS AND METHODS

Patients and tumors

A total of 18 primary breast tumors (14 infiltrating ductal and four infiltrating lobular carcinomas obtained from 17 different patients; one patient had two primary tumors diagnosed at different times) were analyzed in this study. Tissue samples were quickly frozen in liquid nitrogen and stored at −80°C. Ten patients were lymph node positive and only five patients were negative for both estrogen (ER) and progesterone receptors (PgR). ER and PgR status was determined by immunohistochemistry.

RNA extraction from cells and breast tumors

The breast cancer cell lines ZR75.1, SKBR3, T47D, MDAMB231 and MCF7 were harvested by scraping and mRNA was prepared using the acidic method (Chomczynski-Sacchi) (10). All frozen tumors were placed in a lysis solution (guanidine thiocyanate 4 M, sodium citrate 25 mM, sodium N-lauroylsarcosinate 0.5%, β-mercaptoethanol 0.1 M) and homogenized using a Mikro-dismembrator (B. Braun, Melsungen, Germany). Total RNA from biopsies was isolated with the acidic method previously described. Total RNA was pretreated with DNase I (Promega, Madison, WI) before retrotranscription. Reverse transcription reactions were done in 20 µL containing 1 µg of total RNA, 1 mM of each deoxynucleotide triphosphate (Promega), 200 units of MMLV reverse transcriptase enzyme (Promega), 2.5 µM of random primers (Promega), 5X RT-Buffer (Promega) and 40 units of ribonuclease inhibitor (Promega). GAPDH amplification by RT-PCR of all samples was performed to ascertain that equal amounts of each sample were available.

Real-time RT-PCR

Eight genes, not previously described as being estrogen-responsive, were selected out of the set of 344 estrogen-regulated genes found by microarray analysis (9): BCL2-antagonist/killer 1 (BAK-1) (AI741331); retinoid X receptor alpha (RXRα) (AA865269); CD24 antigen (small cell lung carcinoma cluster 4 antigen) (CD24) (AI310309); interferon alpha-inducible protein (clone IFI-6-16) (G1P3) (U22970); TGFB-inducible early growth response gene (TIEG) (AF050110); CD44 antigen (homing function and Indian blood group system) (CD44) (X55150); histone acetyltransferase 1 (HAT-1) (AF030424) and neuropilin-1 (NRP-1) (AF016050).

Real-time, quantitative RT-PCR analyses (11) were performed with an ABI Prism 7700 Sequence Detector and reagents were provided by the same manufacturer (Applied Biosystems). The cDNA sequences present in the arrays were obtained from Incyte Genomics and checked by comparing them with the corresponding GenBank sequence files. Specific primers and probes for the mRNAs to be analyzed were designed using the dedicated software Primer Express 1.0 (Applied Biosystems). Amplicons of 64-101 bps were selected for testing. The melting temperatures of the primers were 59-61°C and those of the probes 69-71°C. Primer and probe sequences are available on request.

RT-PCR reactions were performed using amounts of cDNA corresponding to 0.025-0.125 µg of DNA-treated total RNA, as recommended by the manufacturer. Each 25 µL reaction mix also contained 2X Universal Master Mix (Applied Biosystems), 100 nM TaqMan probe, 20X PDER 18S (Pre-Developed TaqMan Assay Reagent, Applied Biosystems), 300 nM of each primer for AI310309, U22970, AF050110, X55150 and AF030424; 300/600 nM for AF016050 and AI741331; 600/300 nM for AA865269. After 2 min at 50°C to allow AmpErase uracil-N-glycosylase (UNG) to destroy potential contaminant PCR products and 10 min at 95°C to denature UNG and activate Taq polymerase, amplification was carried out in 40 cycles, each lasting 15 sec at 95°C and 1 min at 60°C. A dissociation curve was generated at the end of the PCR cycle to verify that a single product was amplified.

The threshold cycle, Ct, which correlates inversely with the target cDNA levels, was measured as the cycle...
number at which the reporter fluorescent emission increased above a threshold level. To quantitate the target cDNA relative to the zero point, the Ct of each sample was subtracted from the Ct0, and each was corrected for the Ct of the 18S ribosomal RNA (Ct) used as internal reference and coamplified with the target, as described above. The relative quantity of the template (Q) was calculated as: \[ Q = 2^{(Ct0-Ct)-(Ct0r-Ct)} \]

The relative expression values were correlated with a number of clinicopathological factors: the mean rank values for each gene in groups of tumors defined by clinicopathological parameters are shown in Table I. Tumor size, grading, lymph node invasion and p53, ErbB2 and ER immunohistochemical expression were used to stratify the samples. The nonparametric two independent samples tests procedure was used to compare two groups of cases on one variable and the Mann-Whitney U statistic and Wilcoxon statistic were applied to test for significance. A statistically significant positive correlation was found between tumor size (T) and CD24, while a statistically significant negative correlation was found between N and CD44, ER and BAK-1, and ER and HAT-1.

The distributions of observed values for each gene, expressed as the log2 of the relative expression level, are shown in Figure 1. For each gene, the median, interquartile range (25-75%), outliers and extreme cases of individual variables are given. With the exception of RXRα, none of the distributions were normal, and there were occasional erratic outliers.

Among the genes repressed by estrogen in the experimental model system, BAK-1 and G1P3 had median values below the reference control as well as 10 out of 18 and 11 out of 18 of the individual values, respectively, while the CD24 interquartile range was located above the control. In the group of experimentally induced genes, all the interquartile ranges were placed at higher positions on the scale, with medians around (HAT 1) or above (CD44, NRP-1 and TIEG) the reference control. Furthermore, there were more outliers and extreme cases in this group. This behavior may suggest a correlation with the results observed in the in vitro model.

Gene expression profiling divides tumors into high-positive and low-positive to ER

Since all the genes examined were regulated by estrogen in vitro, it was predicted that their status in tumors would correlate with the presence of estrogen receptors, or, more in general, with hormone responsiveness of the tumor. To confirm this hypothesis, the data were submit-
Estrogen-regulated genes analysis in breast cancer

Unsupervised hierarchical clustering analysis was performed and a two-cluster solution was selected (Fig. 2). Fast color rendering of expression data is shown in Figure 3. Clinicopathological factors were investigated in the two clusters using ANOVA for independent samples. A significant statistical correlation was observed with ER status, expressed as the percentage of positive cells by immunohistochemistry (Tab. II). A mean value of 23.43 was found in the “low-positive” cluster and a mean value of 67.36 in the “high-positive” cluster (p=0.02).

Five genes are relevant to discriminate ER status

Multivariate regression analysis was performed to discriminate the most significant variables among the available factors. A backward stepwise method was applied, introducing as covariates the eight genes and clinicopathological parameters. Four of the eight genes together with ErbB2 proved able to predict the ER status.
with an acceptable approximation (R²=0.8). These genes were TIEG, NRP-1, HAT-1 (induced by estrogen) and CD24 (repressed by estrogen).

DISCUSSION

The presence of ER in tumors is a weak marker of the sensitivity of tumor cells to estrogens and, consequently, antiestrogens. The response rate to antiestrogen treatment in patients with advanced ER+ breast cancer is about 30%. The concurrent presence of PgR enhances the chances of response to 70%. This can be explained by the fact that PgR is an estrogen-responsive gene that “marks” the actual functioning of ER in tumor cells.

It is likely that the expression levels of other estrogen-responsive genes will increase the possibility of predicting tumor responsiveness. Genes that are identified through the screening of in vitro model systems treated with estrogen should be evaluated in tumors to determine their expression level, their correlation with receptor status and other clinicopathological variables, and their association with response to therapy. In this work we examined the expression of eight genes that were identified by such screening in a small group of breast tumors; we aimed to find out if these genes were able to discriminate hormone responsiveness by examining their association with receptor status.

At least one report on the expression of each of these genes except G1P3 and HAT-1 in breast cancer was found in the literature. However, none of these studies investigated the direct association with response to endocrine therapy or the correlation with hormone receptor status in tumor biopsies.

CD24 is a surface glycoprotein whose involvement in the metastatic process has been suggested, either alone or in combination with CD44, a transmembrane glycoprotein (12-14). Noteworthy, CD44 was associated with lymph node invasion in our study. Furthermore, in our experimental model CD24 was repressed by E2 treatment while CD44 was induced. Interestingly, it was recently reported that only the CD44+/CD24–/low tumor cell lineage subpopulation, among the breast cancer cells implanted after surgery and grown in immunocompromised mice, proved to be able to form new tumors in mice (15).

NRP-1 is another multifunctional and multidomain transmembrane glycoprotein abundantly expressed in several tumor cell lines including those derived from breast cancer, which acts as a negative regulator of tumor cell apoptosis (16). NRP-1 was expressed at high levels in our tumor samples and was negatively correlated with BAK-1, a proapoptotic regulator. The regulation of BAK-1 was examined in the human breast cancer cell line MCF-7 (17). Treatment of MCF-7 cells with 17β-estradiol resulted in inhibition of BAK-1 mRNA and protein, while progesterone did not affect BAK-1 expression. This means that exposure to estradiol creates an antiapoptotic environment through upregulation of Bcl-2 and downregulation of BAK-1, increasing the survival probability of mammary epithelial cells. Interestingly, our results show a negative correlation between BAK-1 mRNA level and ER status. This also explains how antiestrogens can induce apoptosis in tumor cells.

RXRa is a nuclear receptor essential for the response to retinoids, which also works as a common coreceptor for all nuclear receptor proteins. Higher expression of RXRa was seen in breast cancer compared with benign breast tissue (18). Retinoid receptors modulate various effects of retinoids, including estrogen metabolism in human breast carcinomas. The use of synthetic receptor ag-

TABLE II - MEAN ER VALUES IN THE TWO SELECTED CLUSTERS

<table>
<thead>
<tr>
<th>Group</th>
<th>No.</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>23.43</td>
<td>28.19</td>
<td>0.021</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>67.36</td>
<td>39.13</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>50.28</td>
<td>40.83</td>
<td></td>
</tr>
</tbody>
</table>

Values are expressed as percentages of positive cells by immunohistochemistry.
Estrogen-regulated genes analysis in breast cancer

onists and antagonists clearly showed that the effect of the retinoid on the breakdown of retinoid receptors is receptor-ligand agonist-dependent and blunted by antagonist (19). One study (20) suggests that the receptor-selective retinoid LDG1069 is able to activate retinoid X receptors, to inhibit T47D growth in vitro, and to suppress mammary tumorigenesis in transgenic mice without toxicity. These kind of agents could be considered promising for the prevention of breast cancer and they may be particularly useful in preventing ER-negative breast cancer (21). It was also demonstrated that RXRα is involved in the regulation of aromatase activity in MCF-7 cells, where it mediates the inhibition of aromatase expression in healthy and diseased breast tissue, so some compounds could be useful in the treatment of estrogen-dependent breast cancers (22). RXRα was clearly downregulated after estrogen treatment in our experimental system, but no correlation with ER status was observed in tumors.

TIEG is a transcription factor-encoding gene, characterized as a gene induced early in response to TGF-β which decreases with advancing stages of breast cancer. One study describes TIEG as a gene early induced by 17β-estradiol (E2) in ER-positive human fetal osteoblastic (hFOB/ER) cells, whose effect is E2-induced inhibition of DNA synthesis (19). TIEG expression was found to be high in normal breast tissue, lower in in situ carcinoma, and absent in invasive tumor cells (20). Conversely, TIEG mRNA expression was very high in our tumor samples. It should be noted, however, that our measure is not absolute but relative to a pool of cultured cells. It is therefore likely that higher levels found in tumor biopsies than in cultured cells are simply due to the presence in biopsies of variable amounts of normal epithelium and in situ carcinoma components, whereas all cell lines in the reference pool were derived from highly metastatic tumors.

HAT-1 is a histone acetyltransferase whose expression level was found to be directly correlated with ER status in this study. Since histone acetylation and transcriptional activation of many genes occur after estrogen stimulation, increased levels of this enzyme may play a role in estrogen response. Finally, G1P3 is an immediate response interferon-stimulated gene; the type I (α and β) interferons induce transcription of this gene (21), which has been postulated to encode the precursor of a membrane protein.

In conclusion, with the exception of BAK-1 and HAT-1, no significant correlation was demonstrated between the expression level of the individual genes and ER status in tumor biopsies, in spite of their proven regulation in the experimental model system (9). A very likely explanation for this phenomenon is that many different regulatory pathways apart from estrogen may play a role in transcriptional regulation, so that the contribution of ER to the final steady-state expression of each gene may be masked or undetectable. Combinatorial analysis of several low-penetrance markers is an interesting approach in this respect. In fact, the advances in analytical molecular biology techniques have given us the opportunity to analyze many different markers in small amounts of biological tissue. For example, combinatorial analysis of genomewide gene expression results was demonstrated to be extremely efficient in predicting survival in breast cancer, without the contribution of lymph node invasion (8).

We used unsupervised hierarchical cluster analysis to show the combinatorial ability of the analyzed genes in detecting the ER+ phenotype. The algorithm used was able to cluster the expression profiles into two groups that were significantly different with respect to ER status, thus demonstrating that the combination of several markers identified in vitro by hormone treatments can be a useful tool to discriminate estrogen-responsive tumors. The predictive pool was refined using a backward stepwise method to show that the expression level of only four genes, together with the ErbB2 value obtained by immunohistochemical analysis, can identify the ER phenotype in tumors.

The present work provides data on the expression of eight genes relatively unexplored in breast cancer. It is suggested that combinatorial data analysis may unmask the value of markers which by themselves are not sufficiently informative. These data represent a proof of concept that extending this kind of analysis to the complete set of estrogen-regulated genes should provide an “absolute” profile for the prediction of response to endocrine treatments.

ACKNOWLEDGEMENTS

Supported by grants from the Associazione Italiana per la Ricerca sul Cancro: Finanziamento 2002 (M.d.B. and A.W.), Ministero di Istruzione, Università e Ricerca: Grant FIRB No. RBNE0157EH (M.d.B. and A.W.) and Grant PRIN 2002-03 No. 2002067514 (A.W.), Regione Piemonte and Seconda Università degli Studi di Napoli: Fondi ricerca di Ateneo 2002 (A.W.).

Address for correspondence:
Prof. Michele De Bortoli
Institute for Cancer Research and Treatment (IRCC)
Laboratory of Oncological Gynecology
S.da Provinciale 142, km 3.95
10060 Candiolo (TO), Italy
e-mail: michele.debortoli@unito.it
REFERENCES


