Gene Expression Profiling of Breast Cancer in Relation to Estrogen Receptor Status and Estrogen-Metabolizing Enzymes: Clinical Implications

Vessela N. Kristensen,1 Therese Sorlie,1 Jurgen Geisler,3 Anita Langerød,1 Nobuhiro Harada,4 Rolf Kåresen,2 Noriko Yoshimura,4 P.E. Lønning,3 and Anne-Lise Borresen-Dale1

1Department of Genetics, Institute of Cancer Research, Norwegian Radium Hospital; 2Department of Surgery, Ullevaal Hospital, Oslo, Norway; 3Department of Oncology, Haukeland Hospital, Bergen, Norway; and 4Department of Biochemistry, School of Medicine, Fujita Health University, Toyoake, Japan

ABSTRACT

Interactions between luminal epithelial cells and their surrounding microenvironment govern the normal development and function of the mammary gland. Estradiol plays a key role in abnormal intracellular signaling, which contributes to the development and progression of breast tumors. The present article summarizes the results from a microarray whole genome gene expression analysis as well as a quantitative analysis of the mRNA expression of members of the estradiol metabolic and signaling pathways in the tumors of postmenopausal breast cancer patients. The analysis of the variation in whole genome gene expression resulted in a tumor classification comprising several distinct groups with distinct expression of the estrogen receptor (ER). The parallel study on the expression of only nine mRNA transcripts of members of the estradiol pathways resulted in two main clusters, representing ER− and ER+ tumors. The mRNA expression of the estradiol-metabolizing enzymes did not follow the expression of the ER in all cases, leading to the recognition of several further subclasses of tumors. When the tumor classes obtained by whole genome gene expression analysis were compared with those obtained by independent quantitation of the estradiol-metabolizing enzymes, a statistically significant association between both classification groups was observed. These findings point to a possible association between development of a tumor with a particular expression profile and its capacity to synthesize estradiol as measured by the expression of the transcripts for the necessary key enzymes. Further, whole genome expression patterns were studied in 12 patients treated with anastrozole. Using significance analysis of microarrays, we identified 298 genes significantly differently expressed between partial response and progressive disease groups.

INTRODUCTION

The normal ductal-lobular system of the breast is lined by two epithelial cell types, inner luminal secretory cells and outer contractile myoepithelial cells. Signaling between luminal epithelial cells and their surrounding microenvironment governs the normal development and function of the mammary gland. Variations in transcriptional programs in the cells of breast tumors may reflect these interactions and may bring insight into tumor initiation. The estrogen receptor (ER) plays a central role in the cross-talk between different signaling pathways. The activity of the ER itself is dependent on the synthesis and availability of its ligand, estradiol. Control of this process should be important in breast cancer promotion and may be reflected in the tumor classes observed by gene expression profiling.

With this in mind, we compared the tumor classes obtained by whole genome gene expression analysis (1) with those obtained by previously published independent quantitation of the estradiol-metabolizing enzymes (2): aromatase (CYP19), HSD1, HSD2, STS, and EST as well as the related signaling members ERα, ERβ, cyclin D, and ErbB2 using β-actin as an internal standard.

SUMMARY OF MATERIALS AND METHODS

Patient Material. Results from previously published microarray data (1, 3–5) are summarized in this review. Large-scale microarray data were available for the following patient series:

1. A total of 115 breast cancer patients admitted to the Haukeland Hospital, Bergen, Norway (3–5). The majority of the patients had stage III and IV disease, with a mean age at diagnosis of 64 years (range, 32-85 years). These patients received doxorubicin or 5-fluorouracil/mitomycin C. All patients with an ERα+ tumor were subsequently treated with tamoxifen for 5 years or to time of relapse.

2. Samples (n = 73) from a consecutive series of breast cancer patients admitted to the City Hospital of Oslo (Ullevaal; mean age at diagnosis, 65 years; range, 28-91 years), with mainly stage 1 and 2 disease (1). Quantitative mRNA expression of genes coding for estradiol-metabolizing enzymes (2) was known for 64 of these 73 cases. Patients received standard treatment with tamoxifen.

3. Samples from 12 postmenopausal patients (mean age, 67 years) suffering from locally advanced, ERα+ tumors treated with anastrozole (1 mg/d p.o.) for 15 weeks before surgery (6).
Microarray expression profiles were studied in the pretreatment biopsies and in samples obtained at final surgery. The primary data tables and the image files are stored in the Stanford Microarray Database (http://genome-www5.stanford.edu/).

**Absolute Quantification of mRNA.** Fluorometrical quantitation was done to determine the absolute mRNA contents using fluorescent dye–labeled primers in the presence of an internal standard for each gene: CYP19, HSD1, HSD2, ERα, ERβ, STS, EST (SULT1E1), cyclin D, ErbB2, and β-actin as described in ref. 2.

**Statistical Analysis.** Clinical variables as well as mRNA levels were collected in a Statistical Package for Science version 11.5 (http://www.spss.org/) file where the initial descriptive statistics, cross-tab (χ² test), and Pearson correlation analysis were done. Average-linkage hierarchical clustering was applied by using the CLUSTER program, and the results were displayed by using TREEVIEW (software available at http://genome-www5.stanford.edu/resources/restech.shtml). Absolute mRNA levels were calculated as relative to median, and clustering analysis of levels above and below the median was done using EPCLUST (http://ep.ebi.ac.uk/EP/EPCLUST/). Gene expression clusters were categorized and the distribution of the genotypes in these categories was studied using χ² statistics.

We applied a recently described analytic method called significance analysis of microarrays to search for genes that correlated with treatment response to anastrozole (7). Class prediction was done by using prediction analysis of microarrays and its capacity to synthesize estradiol as measured by the expression of the transcripts for the necessary key enzymes.

**RESULTS**

**Whole Genome Patterns**

Variation in gene expression patterns was initially characterized in a set of 65 surgical specimens of human breast tumors from 42 different individuals using cDNA microarrays representing 8,102 human genes. Pairs of consecutive samples taken from the same tumor separated by 15 weeks of neoadjuvant treatment revealed distinctive molecular patterns in each tumor. Based on variations in gene expression, the tumors could be classified into an ER− basal epithelial-like group, an ErbB2-overexpressing group, a normal breast-like group, and an ER+ luminal epithelial group (3). After expanding the sample numbers to 73, this last group could be divided into at least two subgroups (luminal epithelial A− and B), each with a distinctive expression profile (4). These tumor subtypes were proven to be reasonably robust by clustering using two different gene sets: (a) a set of 456 cDNA clones previously selected to reflect intrinsic properties of the tumors and (b) a gene set that correlated highly with patient outcome (4). Survival analyses on a subcohort of patients with locally advanced breast cancer uniformly treated in a prospective study (9, 10) showed significantly worse prognosis for the ER− basal-like subtype and a significant difference in outcome for the two ER+ groups. After further expanding the set to 115 malignant breast tumors, the same main groups remained: one basal-like, one ErbB2-overexpressing, two luminal-like, and one normal breast tissue-like group (5). The 534 “intrinsic” genes used in this study helped to subdivide the tumors into classes based on their similar expression levels. Further cluster analyses of two published, independent data sets representing different patient cohorts from different laboratories as well as an additional set of 73 unselected cases analyzed in our laboratory (1), resulted in largely the same breast cancer subtypes.

**Estradiol-Metabolizing Enzymes and Correlation with Whole Genome Patterns**

Of the 73 tumors from series 2 (see SUMMARY OF MATERIALS AND METHODS) with known whole genome expression pattern, 64 were also characterized for mRNA expression of several estradiol-metabolizing enzymes (2) and classified into four groups: I: ERα−ERβ+CYP19+, II: ERα+ERβ−CYP19−, III: ERα−ERβ−CYP19+, and IV: ERα−ERβ−CYP19−. When we compared the classification of these tumors by whole genome expression profiling to classification by these four groups, a statistically significant association was observed (Fig. 1). In part, this association could be easily explained through the expression of ERα: The ERα+ERβ+CYP19+ and the ERα+ERβ−CYP19− were most often of the luminal A− and B type, and the ERα−ERβ−CYP19− group was basal like. However, there were patients in the normal like and the ErbB2/HER-2 subgroup whose tumor, although ERα+, did not express estradiol-metabolizing enzymes. The normal-like group was particularly marked by tumor samples not expressing the ER but still producing estradiol III: ERα−ERβ−CYP19+ (Fig. 1). These findings point to a possible association between development of a tumor with a particular expression profile and its capacity to synthesize estradiol as measured by the expression of the transcripts for the necessary key enzymes.

**Gene Expression Patterns in Patients Receiving Anastrozole**

A whole genome gene expression analysis of 12 breast cancer patients treated with anastrozole revealed a tumor classification similar to that of our previous studies. The tumors with low expression of ERα clustered together and were characterized by a strong basal-like signature highly expressing keratin 5/17, cadherin 3, frizzled, apolipoprotein D, etc. The luminal epithelial tumor cluster, on the other hand, highly expressed ERα, GATA binding protein 3, and N-acetyltransferase. An evident ErbB2 cluster was observed due to the marked overexpression of the ErbB2 gene, GRB7 and PPARβ in this patient material. Further, we wanted to investigate the expression patterns of genes that had been shown previously to be up-regulated and down-regulated by estradiol in MCF-7 cell lines (11) in the tumors of the patients treated with anastrozole, when estradiol synthesis should be inhibited in the tissue. Interestingly, the tumors again segregated in two main clusters, largely representing the ERα+ versus the ERα− cases, without actually having the ERα itself among the genes in the cluster (Fig. 2). This classification was mainly due to other luminal class markers such as GATA3 and TFF1 (trefoil factor 1 breast cancer). Using significance analysis of microarrays, we identified 298 genes significantly differently expressed between the partial response and the progressive disease groups. When we clustered all the
samples with the 298 gene list, most of the tumor pairs clustered together, indicating some “intrinsic” response signature that does not change with the treatment as such (data not shown).

**DISCUSSION**

Although it is becoming evident that mRNA expression tumor classification from independent sample sets reflects some common biology, the driving force behind this observed segregation of tumors is less apparent. The data presented in this article suggest a possible association between development of a tumor with a particular expression profile based on whole genome expression and its capacity to synthesize estradiol as measured by the expression of the transcripts for the necessary key enzymes. Mutations of the breast cancer susceptibility gene BRCA1 have been found to predispose to the basal tumor subtype (5, 12), whereas normal BRCA1 functions are consistent with tumor suppressor activity (13–15). Recent studies indicate that BRCA1 interacts with and regulates the activity of ERα and the androgen receptor (16). Furthermore, ER+ cell lines, such as MCF-7 and ZR-75-1, showed a rapid and sustained activation of extracellular signal-related kinase in response to estradiol that was substantially prevented by wild-type but not mutant BRCA1, providing a link between BRCA1 mutation status (and the following tumor classification) and estradiol-induced signaling (17).

ERα, whose expression is one of the more important clinical variables of breast cancer, remains a key player behind the gene expression separation of the luminal and basal types of tumor classes. A possibility may be that the observed tumor classes represent separate cell types from which these tumors originate. This hypothesis has been addressed in a study comparing the expression patterns of two epithelial cell types, inner luminal secretory cells and outer contractile myoepithelial cells (18). Comprehensive expression profiles of the two normal cell types, using immunomagnetic cell separation and gene...
expression microarray analysis, revealed distinct expression profiles of the two cellular types in breast cancer patients. Whereas between 4% and 20% of the luminal epithelial cells are positive for ER and ~50% are positive for PR, myoepithelial cells do not express ER. Regulation of the normal mammary gland function is closely related to the preservation of the tissue architecture (19). In vivo, myoepithelial and most luminal cells are in contact with the basement membrane, and communication between the luminal cells either directly or through the myoepithelial cells is essential for mammary differentiation. The proteins that form the intermediate filaments of epithelial cells, the cytokeratins, are heterodimers with very distinct expression patterns. High expression of keratin 5 and 17 was observed in the basal-like tumors characterized by ER negativity and BRCA1 mutations (4, 5). The basement membrane molecules laminin-1 and collagen IV were found to be involved in the maintenance of ERx expression in cultured nonmalignant mammary epithelial cells, and this response could be interfered with by disrupting cell/extracellular matrix adhesion (19).

The relationships between the ER and both its ligand and the enzymes that synthesize it are not well understood. We have reported previously a SNP in the 3’untranslated area of CYP19 mRNA to be associated with higher mRNAs levels and increased tumor size in locally advanced breast cancers (stage III; ref. 20). Patients with stage I and II disease with no expression of CYP19, on the other hand, had short disease-free survival (P > 0.0005), possibly attributable to hormone independence and poor response to standard treatment with tamoxifen (2). We have previously reported high correlation of the expression of CYP19 (aromatase) mRNA and a switch from the adipose tissue–specific promoter to the ovary-specific promoter (20). Furthermore, the alternative usage of exon 1 was also associated with shorter disease-free survival; however, the patients with worse prognosis had no expression of CYP19 mRNA at all (2). This is a somewhat paradoxical result, because these and our previous data (20) suggest that in the course of the disease a switch of promoter occurs and results in higher expression levels. The process, however, seems to occur in a given timeframe of the disease on the way to hormone-independent breast cancer. The poor prognosis of patients with null expression of aromatase may also reflect treatment failure of the patients. This was a consecutive series of patients with predominantly stage I/II disease who had been given a standard treatment with tamoxifen. Patients who relapsed after adjuvant tamoxifen then received aromatase inhibitors. Unfortunately, a substantial number of tumors are intrinsically tamoxifen resistant, despite ER positivity, and eventually almost all breast carcinomas acquire resistance to tamoxifen (21). We have no exhaustive data on treatment response to verify whether patients with null aromatase expression in their tumors respond poorly to such standard therapeutic procedure.

Anastrozole has been shown to inhibit both total body and intratumoral aromatization in vivo (6, 22) and to cause tumor regression in the neoadjuvant setting (23). In a recent study (6), we determined the effect of primary medical treatment with anastrozole for 15 weeks on tissue estrogen levels in 12 patients with locally advanced breast cancer. Using a combined high-performance liquid chromatography-RIA method, we found that intratumoral levels of estrone, estradiol, and estrone sulfate decreased by 89% (confidence interval, 73.2-95.5), 83.4% (confidence interval, 63.2-92.5) and 72.9% (confidence interval, 47.2-86.1), respectively (6). Although all patients presented with ER+ tumors in the routine immunohistochemical analysis, in some the receptor expression was low. The estrogen levels

Fig. 2  Hierarchical clustering of breast cancer tumors using lists of genes up-regulated and down-regulated in MCF-7 cell lines treated with estradiol (11). The clustering was strongly driven by genes such as GATA-3, Trefolfi factor 3, IGFBP3, and cdc6.
dropped in all patients; however, the clinical response varied. The reason for lack of response to hormone manipulation in some breast cancer patients expressing the ERα is poorly understood, and larger studies are needed to validate the significance of these genes for treatment response to anastrozole.

In summary, the analysis of the data sets presented here, studying whole genome expression profiles in both consecutive early-stage breast cancer and selected locally advanced cases, both treated or untreated, suggest common tumor classification groups with different disease-free survival rates. Can this observation cast more light on the origin and on the fate of these breast cancers in terms of prognosis and treatment response? The present tumor classification suggests that there is a more “uneventful” path to develop breast cancer, which is ER+, expresses genes characteristic for the luminal cell type and has a better prognosis in terms of disease-free survival time. On the other hand, there are tumors with more dramatic presentations that are ER−, often with amplification of ERBB2, often harboring TP53 mutations, and expressing genes characteristic of the basal cell type. Carriers of germ line mutations in the BRCA1 gene are predominantly of the second type. Do these two types occur independently from different cell entities of the breast (luminal and basal) comprising different types of disease? The existence of intermediate types (e.g., luminal B), which express ER but may also have amplification of ERBB2, express some basal cell markers, may argue against this hypothesis. It is equally feasible to hypothesize that in the course of tumor development cells express markers that make them resemble either the one or the other cell type. The question would then be whether these expression programs occur simultaneously or in succession during the progression of the disease and whether the results we observe at the time point of our microarray analysis are a product of selection and if so what the driving forces of this selection are.

To be able to use microarray data for prognosis and prediction of treatment response and clinical outcome, we need to further dissect the observed clusters of genes in terms of relevant regulatory networks and incorporate hypotheses regarding functional pathways (24). We are currently performing analysis for common regulatory blocks in the separate gene clusters as well as comparing the patient data with data from cell lines subjected to similar treatments.

OPEN DISCUSSION

Dr. Jose Russo: In what specific cluster do you find proliferation markers like Ki-67 or PCNA? Do you find them in one specific cell type, or are they broadly distributed between the two populations?

Dr. Vessela Kristensen: The proliferation genes form a nice cluster by themselves, and we can see that in both the ER+ and mostly in the ER− tumors, that is, in the luminal and the basal types. It is exactly the presence of the proliferation cluster that distinguishes the ER-positive tumors with poor prognosis (luminal type B) from those with good prognosis (luminal type A).

Dr. Kent Osborne: You said that ERβ was associated with a poor prognosis in tamoxifen-treated patients. ERβ has been really difficult to study for lack of an antibody that is reproducible and specific and that works in an immunohistochemistry. We just submitted for publication a study looking at ERβ protein expression in a series of adjuvant tamoxifen-treated patients. There, ERβ was actually a good factor—in fact, independently and statistically significantly so. Some of the data on ERβ function are beginning to show that it puts the brake on the system, to put it in a simplistic fashion.

Dr. Donald McDonnell: In general, ERα activates and ERβ modulates. What the specific roles of these two proteins are depends on what model systems you’ve used. There is an interesting possibility here that you’re looking at mRNA. There is a negative feedback of the protein to the gene, and if the feedback is disrupted, that is, the protein is not being produced, the mRNA is going to compensate and go up. What could happen is that you’re getting a high level of ERβ mRNA in those tissues that are actually not expressing very much ERβ protein.

Dr. Kristensen: Definitely. We observed the same with aromatase. It was in the patients with successful inhibition of the protein that an increase in mRNA levels was observed. Based on mRNA alone, we would have concluded quite wrongly.

Dr. Russo: In those cases where you compare nontreated tumor versus treated, for example, with tamoxifen, is the genomic profile of the tamoxifen-responsive tumor similar to the normal breast or is it a different profile?

Dr. Kristensen: None of these studies are actually ready to answer the question of the treatment response. There are a lot of statistical issues in that, and it is very difficult to separate the important from the unimportant factors.

Dr. Per Lønning: It is interesting, however, to compare Therese Sørlie’s PNAS paper [Proc Natl Acad Sci USA 2003;100:8418–23] where she applied the same gene analysis to the Dutch van’t Veer material. That article reported that the discrimination regarding relapse versus survival was much smaller between the luminal A and B class in the Dutch material compared to the Norwegian material. Now, the difference is that, in the Norwegian material, all of the patients who were ER+ had been exposed to tamoxifen, and in these samples, there was a huge difference between luminal A and luminal B. Whereas in the Dutch material none of these patients received tamoxifen, and the discrimination between luminal A and luminal B regarding relapse was much smaller. So, quite indirectly, that comparison indicates that the luminal B subtype are poor responders to tamoxifen compared to luminal A, but of course this needs to be confirmed.

Dr. Stephen Johnston: The group that would be really interesting to study, based on those data sets, is the patients treated with up-front endocrine therapy in a locally advanced breast cancer setting or a preoperative neoadjuvant setting. Because if you can really differentiate responders from non-responders, particularly in ER+, then that would help us tremendously.

Dr. Richard Santen: In your data, it looked as if expression of aromatase was a negative survival factor. We did some studies years ago that looked at this. The tumors that were highly differentiated (grade 1) had high aromatase activity. As they became less differentiated, the aromatase activity went down, but then when you got to the grade 3 very aggressive tumors, the aromatase activity went back up again. It appears that as a tumor de-differentiates it begins to secrete proteins or make proteins that
weren’t there previously. That observation, if one accepts it, offers two ways of thinking about the biology of aromatase. It is making estrogen, but it may also be a marker of a de-differentiated tumor. That might explain why you have poor prognosis with high aromatase activity. But it really says that you can’t interpret the biology simply by the over- or underexpression of a gene; there are other regulatory factors.

**Dr. Kristensen:** That is very interesting to hear because we do studies on genetic polymorphisms and how they influence the observed phenotypes, and we find such patterns as well. We have a polymorphism in the 3’ area of aromatase that correlates with high expression levels, which correlate with large tumor. So, you would say if this polymorphism is present, more aromatase will be produced, allowing the tumor to become locally advanced. But in a parallel study from our own lab, we see that if there is no expression of aromatase, the prognosis is very poor.

**Dr. McDonnell:** In the article from Dennis Sgroi [Proc Natl Acad Sci USA 2003;100:5974-9] using laser capture microdissection to identify factors that predicted for poor prognosis, two or three of the genes that popped up I had never seen before in any of these analyses. I’m wondering is it because the laser capture microdissection looks only at actual epithelial cancerous cells, whereas in these studies here with larger tumors, you are actually looking at other factors from stromal cells or adipocytes?

**Dr. Kristensen:** Yes, there are pros and cons for both approaches. In this case, we have to use statistical tools to dissect all these genes that are not directly involved in the process. Despite the response differences Prof. Lonning points to, the fact that we see these common classes in both treated and untreated tumors and in all different stages, that means that what we observe is some intrinsic signature. For studies of mechanism, laser capture dissection is more clear-cut because one knows what the experimental settings are. But when it comes to extrapolating these data back to phenotypes like overall survival, then it is more problematic to generalize the observations.

**Dr. McDonnell:** I was really amazed that when you looked at what emerged as prognostic factors from these two approaches, they were completely different. There was very little overlap between them.

**Dr. Lonning:** I think we should focus on how can we use microarrays to screen for functional pathways and genes related to functional pathways instead of focusing on prognostic or predictive factors. You don’t get a predictive factor straight out of a microarray without incorporating a functional hypothesis.

**Dr. Osborne:** We are still lacking a clear picture of biology. For all we know, a lot of these genes that we think now are independent of each other might be in a similar overlapping network. We just don’t know.

REFERENCES

Gene Expression Profiling of Breast Cancer in Relation to Estrogen Receptor Status and Estrogen-Metabolizing Enzymes: Clinical Implications

Vessela N. Kristensen, Therese Sørlie, Jurgen Geisler, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/2/878s

Cited articles
This article cites 24 articles, 10 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/11/2/878s.full.html#ref-list-1

Citing articles
This article has been cited by 6 HighWire-hosted articles. Access the articles at:
/content/11/2/878s.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.