Review Article

Gene expression profiling in breast cancer

Belisario A Arango, Celine L Rivera, Stefan Glück

University of Miami/Sylvester Comprehensive Cancer Center, Miami, FL, USA

Received November 30, 2012; Accepted February 17, 2013; Epub March 28, 2013; Published April 8, 2013

Abstract: In recent years, molecular research has translated into remarkable changes of breast cancer diagnostics and therapeutics. Molecular tests such as the 21 gene expression test (Oncotype DX™) and 70 gene microarray test (MammaPrint®) have revolutionized the predictive and prognostic tools in the clinic. By stratifying the risk of recurrence for patients, the tests are able to provide clinicians with more information on the treatment outcomes of using chemotherapy, HER2 targeted therapy or endocrine therapy or the combination of the therapies for patients with particular genetic expressions. However, it is still questionable for clinical applications as some areas remain unclear and that the true benefit still needs prospective evaluation. Such studies are under way and are anxiously awaited. In this paper, the limitation of the molecular tests are discussed. As we are moving towards personalized medicine, molecular profiling will not only result in better outcomes but in a certain proportion of patients, likely will spare unnecessary use of cytotoxic compounds and reduce the cost to the health care systems.

Keywords: Early breast cancer, gene profiling, predictive and prognostic power

Introduction

Breast cancer is the most common female malignancy in the US, the second most common cause of cancer death in women, and the main cause of death in women ages 40-59 [1]. With the advent of improved imaging techniques and the general population’s awareness of this disease, increasing numbers of patients are more frequently diagnosed with small tumors and negative axillary nodes. Since the likelihood of distant recurrence, for early stage hormone receptor positive breast cancer, in patients treated with tamoxifen alone after surgery, is approximately 15% at 10 years and the absolute survival benefit of chemotherapy in this same group ranges from 3-10%, a significant proportion of women might be overtreated [2, 3].

Despite survival advantages achieved using systemic therapy in many women with early-stage breast cancer; there are also significant toxicities, quality of life-related side effects, potential long term side effects and costs associated with such therapy.

It is well recognized that breast cancer represents a heterogeneous group of tumors with varied morphologic and biological features, clinical behavior, and response to therapy. The most widely used treatment guidelines are the St. Gallen in Europe and NCCN in the United States [4, 5]. These guidelines recommend adjuvant systemic therapy for 85-90% of lymph node negative breast cancers. Prognostic variables used to predict tumor behavior and guide the use of systemic therapy include tumor size, lymph node involvement, tumor grade, histologic type, proliferation status, and growth rate as well as at least 3 receptors (Estrogen receptor [ER] progesterone receptor [PR] and human epidermal growth factor receptor 2 [HER2]. Together they represent the main components of commonly used prognostic algorithms, guidelines, and indices [6-8].

However, these traditional prognostic factors are insufficient to reflect the whole clinical and genetic heterogeneity of breast cancer and are unable to tailor treatment to each patient. Furthermore, there is only modest interobserver concordance in the assessment of histologic grade amongst pathologists, varying from 59-65% with a greater degree of agreement for poorly differentiated tumors [9]. Furthermore, measuring ER, PR, and HER2 with immunohis-
Gene expression profiling in breast cancer

tochemistry (IHC) and fluorescence in situ hybridization (FISH) from different sources and laboratories is not always reliable [10-12]. With the standard methods presently available, such as IHC, the rate of false-negative results for ER and PR determination may range between 30-40% and could even be as high as 60-70%. Similarly, unacceptably high false-negative rate of approximately 20% was observed for HER2 determination by IHC in a community-based study [13-15].

Tumor gene expression profiling has emerged as a tool that can provide additional information, about breast cancer biology and behavior, to the traditional prognostic factors. In this review, we present and update the status of clinically applicable gene expression based assays, and we discuss their reproducibility in the care of breast cancer patients.

**Oncotype DX recurrence score**

Oncotype DX (Genomic Health, Inc., CA) is a RT-qPCR based signature that measures the expression of 21 genes (16 cancer related and 5 reference genes). RNA is obtained from formalin-fixed paraffin-embedded tissue samples. The Oncotype assay is recommended for patients with hormone receptor positive and Her-2 not overexpressed axillary lymph node negative early stage breast cancer. The test requires assessment of estrogen receptor and Her-2 status by an alternative method [9].

The Oncotype DX has become the most widely used clinical gene expression assay in the US. The genes in the assay were selected from 250 candidates that were tested for association with survival in a cohort of 447 tumor samples, from the tamoxifen treated and node negative cases of the National Surgical Adjuvant and Breast Project (NSABP) B-20 clinical trial [16]. Oncotype DX was validated in a large cohort of ER positive node negative breast cancer patients treated with tamoxifen, enrolled in the NSABP B-14 study. In this study, the rates of distant recurrence at 10 years were 6.8%, 14.3%, and 30.5% for the low-risk, intermediate, and high-risk groups, respectively [9].

The recurrence score (RS) is a continuous variable, ranging from 0 to 100, and constitutes a measure of the risk of distant relapse within 10 years. The score is an independent prognostic factor for patients with ER positive, node negative breast cancer treated with adjuvant tamoxifen. Patients can be classified into three categories on the basis of RS. Low risk (RS<18), intermediate (RS 18-31), and high (RS>31), which equate with 10 year relapse rates of 7%, 14%, and 30%, respectively. Women in the low risk group do not seem to benefit from adjuvant chemotherapy as shown in the NSABP-B20 analysis that randomly assigned patients to receive CMF chemotherapy with concurrent tamoxifen (TAM) or TAM alone. Only a small subset of tissues was available for analysis (651 samples from 2363 randomized patients). The analysis showed that their distant metastasis free survival is higher than 90% regardless whether or not they received CMF chemotherapy. In contrast, women in the high-risk group derived benefit from adjuvant CMF chemotherapy [17]. The question remains unanswered for women who fall into the intermediate-risk group. The current ongoing Trial Assigning IndividuaLized Options for Treatment (TAILORx) study seeks to answer this question by randomly assigning patients with intermediate RS to adjuvant chemotherapy followed by anti estrogen treatment or anti estrogen treatment alone. This randomized controlled trial finished accrual and first data are expected in late 2014.

**MammaPrint**

MammaPrint (Agendia, Irvine, CA and Amsterdam, The Netherlands) is a microarray-based gene-expression profiling assay of RNA. The test is comprised of 70 genes identified from an initially unselected set of all>25,000 genes within the human genome which were obtained from fresh frozen samples of tumor tissue [18]. New studies have demonstrated that the test could be done by RT-qPCR, both in fresh frozen and formalin-fixed paraffin embedded tissue, with equivalent performance [19]. The genes are associated with all hallmarks of cancer including proliferation, invasion, and angiogenesis. The genes were obtained from tissue of 78 patients with lymph node negative breast cancers, most of which were ER positive tumors and did not receive adjuvant systemic therapy [18]. This signature has been validated on numerous cohorts of node negative patients, and has demonstrated to provide independent prognostic information beyond standard clinicopathological variables such as age, histological grade and pathological stage [18, 20, 21].
The test can be used in ER-negative breast tumors.

Mammaprint is the first and so far only FDA-approved gene-expression assay to be used as prognostic test for women with node-negative breast cancers. The test yields two prognostic groups: low-risk and high-risk. This signature is predictive of both distant disease-free survival and overall survival when adjusted for lymph node status. Patients in the low-risk group have a distant metastasis free survival of over 90% without the addition of systemic chemotherapy. The test also provides predictive information, i.e., women in the high risk group derive a benefit from adjuvant chemotherapy [18]. In a retrospective study of neoadjuvant chemotherapy for patients with locally advanced disease, none of the patients in the low risk group had a complete pathological response to treatment [22]. In a study that classified breast cancers into molecular subtypes and evaluated the response to neoadjuvant chemotherapy and long term outcomes using MammaPrint together with an 80-gene molecular subtyping profile (BluePrint), in 421 patients. The combination of MammaPrint and BluePrint resulted in 4 distinct molecular groups: Luminal A (MammaPrint Low Risk/Luminal-type), Luminal B (MammaPrint High Risk/Luminal-type), Basal-type and HER2-type. Luminal A patients (BluePrint Luminal/MammaPrint Low Risk) have a good baseline prognosis with excellent survival and may have no benefit from chemotherapy. BluePrint classifies more patients as Basal-type (n=120) with higher pCR rate (42%), compared to clinical subtyping (n=93) with a pCR rate of 31%. Identifying that molecular subtyping can improve stratification of patients in the neoadjuvant setting; MammaPrint low risk patients have a good baseline prognosis with excellent survival and may not benefit from chemotherapy. This indicates that Mammaprint helps to establish a clinical correlation between molecular subtyping and treatment outcomes [23].

**Mapquant DX genomic grade index**

Mapquant DX (Ipsogen, Marseilles, France) is a prognostic and predictive signature created with the intention to further stratify tumors based on their histologic grade and hence diminishing the interobserver variability. The test not only distinguishes between grade 1 and grade 3 tumors, but also divides grade 2 into two categories with high versus low risk of recurrence. Mapquant DX downgrades 70% of grade 2 tumors to grade 1 [24]. The test takes into account 97 genes, which are mostly involved in cell cycle regulation and proliferation, to compose a genomic grade index (GGI). Patients in the high risk groups are recommended to receive adjuvant chemotherapy. In patients receiving taxane plus anthracycline chemotherapy, high GGI was associated with excellent response to chemotherapy [24]. The prognostic information provided by the GGI is only applicable to ER-positive tumors.

As Mammaprint, the Mapquant DX GGI initially required tumor tissue in fresh frozen samples, but now a modified version based on RT-qPCR has been developed [25].

**Rotterdam signature**

The Rotterdam 76 gene signature was developed on the basis of supervised analysis of microarray data in a training set of 115 breast cancers. ER positive and ER negative tumors were evaluated separately, leading to the identification of 60 genes in the ER positive tumors and 16 genes in the ER negative tumors. This 76-gene prognostic signature can be used to predict the development of distant metastasis within 5 years in lymphnode negative breast cancer patients, irrespective of age and tumor size, who did not receive systemic chemotherapy [26]. In 198 patients, the 76 gene signature was a strong prognostic factor and outperformed the St Gallen’s and NCI guidelines in identifying patients with good prognosis in which adjuvant chemotherapy was not needed. The median follow-up was 14 years and distant metastasis was observed in 51 (26%) of patients, with 35 (18%) of them showing progression within 5 years. When the patients were compared for high clinical risk by the Adjuvant! Online software 152 (77%) were considered high risk and 46 (23%) low risk. The 76 gene signature identified 143 (72%) high genomic risk and 55 (28%) low genomic risk. Interestingly, the low genomic risk group contained 21.9% (14 of 64) of all ER negative patients, whereas the clinical low risk group did not contain any. The time to distant metastasis at 5 and 10 years were 98% and 94% for the low genomic risk group and 76% and 75% for the high genomic risk group, respectively. The 5
Gene expression profiling in breast cancer

and 10 year overall survival were 98% and 87% for the low genomic risk group and 84% and 72% for the high genomic risk group, respectively. Both were statistically significant with time to distant metastasis HR 5.78 and 2.87 for the OS [27].

PAM50

In a neoadjuvant study of 157 patients with early breast cancer [28] we have shown that tumors with luminal signature have very low pCR and npCR, and those with HER2 enriched and basal subtype, particularly with a p53 mutation have high pCR rates with only 4 cycles of preoperative chemotherapy. Invasive breast cancers can be classified by whole gene arrays into at least four major biologic “intrinsic” subtypes referred to as Luminal A, Luminal B, HER2-enriched, and Basal-like. These subtypes have been reproducibly identified in the research setting by microarray and RT-PCR. In 2009, Parker et al. proposed a 50-gene set, a Prediction Analysis of Microarray (PAM50), for standardizing subtype classification. The PAM50 Breast Cancer Intrinsic Classifier is the clinical manifestation of this gene set using a qRT-PCR assay that has been validated on formalin-fixed paraffin-embedded tissues. The test measures the expression of 50 classifier genes and five control genes to identify the intrinsic subtypes of breast cancer known as Luminal A, Luminal B, HER2-enriched and Basal-like. Multivariable analyses using the PAM50 subtypes and other clinical data (e.g., node status, grade, ER-status) show that the PAM50 is an independent prognostic test of survival in breast cancer [29].

The PAM50 test provides additional information about the tumor biology and quantitative data on biomarkers already used for treatment decisions. Along with a categorical classification of breast cancer subtype, the clinical PAM50 test also provides quantitative values for proliferation, luminal gene expression, ESR1, PGR, and ERBB2. The genes used for subtyping are provided in Figure 1.

Luminal A tumors usually have intermediate to high expression of ESR1 and ER-regulated genes and rarely have high ERBB2 expression. Luminal B tumors usually have intermediate to high expression of ESR1 and estrogen-regulated genes and often have higher proliferation than Luminal A tumors. HER2-enriched tumors usually have intermediate to high expression of the ERBB2 gene and intermediate to low expression of ESR1 and estrogen-regulated genes. Approximately one-third of tumors subtyped as HER2-enriched are not HER2+ by IHC (2+ or 3+ HER2 score) or fluorescence in-situ hybridization (DNA amplified for ERBB2). Basal-like tumors usually have low expression of ESR1, PGR, ERBB2, and estrogen-regulated genes, but have high proliferation.

Discussion

The natural history of breast cancer is changing as the benefits of screening mammography and adjuvant chemotherapy are evident with earlier diagnosis of smaller tumors without lymph node involvement. Thus, the need for better stratification of patients is becoming more important in order to identify those patients who will not need to be treated with adjuvant chemotherapy after optimal locoregional treatment, those who will have endocrine sensitive disease to be treated with anti
Gene expression profiling in breast cancer

estrogen against as well as identifying those high-risk patients who will benefit from systemic chemotherapy.

Gene expression profiling has taken front stage in the past decade and some of these test seemed to have outshined traditional predictive and prognostic factors. However, many problems need to be addressed. No randomized controlled trials are at present reported yet. These assays are performed by specialty laboratories which require cancer tissue to be sent to them. In order for these assays to be more widely used there is the need to develop a standardized assay that can be reproduced by commercial and academic laboratories throughout the world. In addition, the biological roles of the genes included in most of these gene arrays are not completely understood and it is often unclear which clinical or tumor characteristics are being measured. Furthermore, prognostic signatures rely heavily on the prognostic power of proliferation related genes. The prognostic information offered by these assays is limited and the development of gene signatures to predict response to specific agents has been less successful.

The clinical utility of any of these tests has been tested in three randomized prospective trials; 1) TAILORx: is a prospective phase III study has randomized approximately 10,000 women with intermediate risk, by OncotypeDx, early stage hormone receptor positive breast cancer to receive adjuvant chemotherapy followed by antiestrogen or antiestrogen alone. 2) MINDACT (Microarray In Node-negative Disease may Avoid ChemoTherapy): is a multicentre, prospective, phase III randomised study comparing the 70-gene expression signature with a common clinical-pathological prognostic tool (Adjuvant! Online) in selecting patients for adjuvant chemotherapy in node-negative breast cancer. Patients with concordant Mammaprint and Adjuvant Online will be randomized to receive chemotherapy. Patients with discordant Mammaprint and Adjuvant Online will be randomized to receive chemotherapy plus antiestrogen or antiestrogen alone based on the result of either Mammaprint or Adjuvant Online alone. 3) RESPONDER: is a prospective phase III study that randomize women with intermediate risk breast cancer, by OncotypeDx, hormone receptor positive breast cancer and 1 to 3 lymph nodes positive to receive adjuvant chemotherapy followed by antiestrogen or antiestrogen alone. Data will become available only in a few years. For now we should use this technology with caution and only as a complement to traditional biologic and morphologic prognostic and predictive information rather than a replacement of the latter. Traditional pathology will remain important for diagnosis and selection of optimal tumor tissue for microarray analysis [30].

In the near future we hope to have the capacity to also determine sensitivity to specific therapeutic agents. One of such tests based on cDNA microarray is TheraPrint (Agendia). Its clinical utility also is a topic of research and might guide us in the future not only whether or not to use adjuvant therapy but additionally, which compounds might be most effective.

Address correspondence to: Dr. Stefan Glück, Department of Medicine, Division of Hematology/Oncology, Sylvester Comprehensive Cancer Center, University of Miami, Leonard M. Miller School of Medicine, Miami, FL, USA. Phone: 305-243-1542; E-mail: SGluck@med.miami.edu

References

Gene expression profiling in breast cancer


Gene expression profiling in breast cancer


