Gene profile and response to treatment

J. Smeds1, L. D. Miller2, J. Bjöhle1, P. Hall3, S. Klaar1, E. T. Liu2, Y. Pawitan3, A. Ploner3 & J. Bergh1

1Department of Oncology and Pathology, Radiumhemmet, Karolinska Institute and University Hospital, Stockholm, Sweden; 2Genome Institute of Singapore, Singapore; 3Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

Introduction

Breast cancer is the most common malignancy in women, with more than 1 million new individuals diagnosed annually [1]. It is now firmly established that the use of adjuvant chemotherapy and/or endocrine therapies in breast cancer management will result in clinically meaningful reductions in breast cancer recurrence and improved overall survival [2, 3]. Current international guidelines recommend adjuvant systemic therapy for a large fraction of breast cancer patients, including all patients with primary axillary lymph node-positive disease and up to 85–90% of patients with lymph node-negative disease [4–6]. Although the effect of adjuvant therapies will be as much as a 50% relative reduction of disease recurrence, the other half of the patients would have been cured without the use of adjuvant therapies or would have relapsed despite their use. These shortcomings of over- and undertreatment are owing to limitations of the presently accepted prognostic and therapy-predictive factors, which lack sufficient sensitivity and specificity. Better prognostic and therapy-predictive factors (besides oestrogen and progesterone receptors and HER2/neu) are urgently needed.

Presently used prognostic and therapy-predictive factors

The St Gallen guidelines recommend the use of age, tumour size, histopathological grade and axillary lymph node status for prognostication [5]. These factors are claimed to be true prognostic factors, although they have never been investigated in prospective and randomised studies and accordingly they do not fulfil the requirement for evidence level 1. The only accepted therapy-predictive factor in the St Gallen recommendations is receptor status [5]. At the most recent St Gallen consensus conference, vascular invasion and HER2/neu were accepted as additional risk factors for prognostication.

Overexpression/amplification of the oncogene HER2/neu occurs in ~20–30% of primary breast cancers and these patients have a worse prognosis [7, 8]. Overexpression/amplification of HER2/neu has also been associated with a better response to anthracyclines compared with CMF-based therapy, as well as a better response to aromatase inhibitors compared with tamoxifen [9–11]. Furthermore, overexpression of HER2/neu has been described to be associated with a worse response to tamoxifen [12–14], while other studies have failed to demonstrate this relationship [15]. Although these (prognostic and predictive) factors perform quite well using group-based statistical analyses, the outcome for the individual patient contains a rather high degree of uncertainty.

Limitations and problems for prognostication and therapy-predictive factors

Ideally, a prognostic factor should be evaluated in untreated patients. Today, however, the majority of breast cancer patients will receive adjuvant therapy that will often influence outcome, and primary tumour specimens from untreated breast cancer patients (particularly from more recent time periods) are likely to be biased towards those with a relatively low risk for recurrence, or representative of a group with poor outcome owing to other serious medical conditions contra-indicating the use of adjuvant therapies. In addition, almost all breast cancer patients receive therapy for metastatic disease, which also influences survival [16–20]. Thus, the ideal setting for testing prognostic factors is virtually unattainable.

A further concern is that many investigators would argue that prognostic and therapy-predictive factors need to be tested in randomised and prospective studies. A randomised study is of course highly relevant for testing a few pre-defined questions, but randomised studies will frequently comprise a highly selected cohort of patients not representative of the general breast cancer population. For example, the inclusion and exclusion criteria for randomised studies are often very strict. Accordingly, patients with co-morbidities or other malignancies who have an expected shorter survival will automatically be excluded. The randomised patient population will thus very likely be comprised of individuals with better prognosis and outcome than those that are truly unselected and representative of the population. We therefore suggest that a new prognostic factor or therapy-predictive factor should be tested on unselected and population-based materials, with a complete coverage over a predefined time period and geographical region.

Microarray-based profiling for biological understanding and prognosis

Recently, gene expression profiling of primary breast cancers has been used to improve prognostication for primary breast cancer patients. This approach is abnormal, and the outcomes are still relatively less reproducible and reliable compared with the presently accepted prognostic and therapy-predictive factors. With the ongoing development of new microarray-based profiling technologies, these outcomes will likely improve in the future.
Two Dutch studies investigated 295 and 286 primary breast cancers, respectively [6, 22]. The Amsterdam group identified a 70-gene set that separated both node-negative and node-positive patients into good and poor prognosis groups better than the St Gallen and NIH consensus risk criteria [21, 22]. The Rotterdam group identified a 76-gene signature for untreated node-negative patients, which performed better in the multivariate analysis compared with classical breast cancer prognostic factors [6]. Sixteen genes were needed for the oestrogen receptor (ERα)-negative group, while 60 genes were needed for the ERα-positive group, further underlining the importance of ERα status in breast cancer biology [6]. Though only three genes overlap between the expression signatures uncovered by these two groups, both signatures contain genes belonging to same molecular pathways, such as proliferation and apoptosis, and are similarly effective in prognosis. Thus, they both have the potential to guide the reduced use of adjuvant chemotherapy for node-negative patients.

In Europe, a prospective and randomised study called MIN- DACT has been planned that seeks to validate the Amsterdam gene signature [25], and the signature has been tested in a preliminary analysis of ~300 node-negative frozen samples from six different institutions, which resulted in a somewhat lower prognostic performance [25].

We, also, have recently analysed untreated and treated primary breast cancers using the Affymetrix U133A/B microarray platform for more than 400 patients [26]. The prognostic gene signature that we derived was similar in gene number to the Dutch signatures, but only three genes overlapped with the Amsterdam signature and a single gene with the Rotterdam signature [6, 21]. That several similar studies consistently identify prognostic gene signatures of variable composition might be explained by a number of contributing factors, including: (i) different microarray platforms; (ii) cohort biases owing to dissimilar patient selection procedures and sample size; and (iii) different normalisation approaches and statistical methods for gene selection.

Microarray profiling reveals that breast cancer consists of many diseases

Microarray profiling has also been used to discover distinct molecular subgroups of breast cancer that are not discernable by morphological classification [27–29]. For example, the clinical entity of breast cancer can be separated into five or six distinct subgroups with statistically significant prognostic separations both for relapse-free and overall survival [28, 30]. The presumed important breast cancer-related genes p53, HER2/neu and ERα were observed to present in high frequency in certain of these microarray-based subtypes of breast cancer [28]. Five of seven and nine of 11 of the tumours in the HER2/neu and basal cell-like subclasses, respectively, had p53 mutations [28]. In the large luminal subtype A class, only 13% of the tumours contained a p53 mutation (Figure 1), consistent with this group’s significantly better outcome. The ERα gene on the other hand, was expressed in highest levels in the luminal subtype A. The same research consortium from Norway and USA expanded their initial study of 78 malignant breast tumours to a total of 115 breast cancers and subdivided them into the basal-, HER2/neu overexpressing-, luminal A- and B- and normal-like subtypes [30]. Excluding the normal-like subtype, they confirmed previous observations and validated the prognostic implications for the other subtypes using the publicly available gene list from the Amsterdam group consisting of 78 breast cancers [21, 30]. In this study, the consortium also studied 18 and two tumours from patients with hereditary breast cancers with BRCA1 and BRCA2 mutations, respectively. The BRCA1-mutated tumours clustered to the basal-like type, consistent with the poor prognosis associated with BRCA1 mutants [30]. Together, these findings are intriguing in that they suggest that breast tumour subclasses, defined by distinct molecular configurations, may have prognostic primacy over conventional single-gene factors.

Gene expression profiling has also been used to discriminate tumour classes associated with known biomarkers of high pathological and therapeutic relevance. In one study, the molecular profiles of spontaneous breast cancers were compared with those with BRCA1 and BRCA2 mutations [31]. Almost 200 genes were found to be significantly differently expressed between tumours with BRCA1 and BRCA2 mutations [31]. Interestingly, one patient with a sporadic breast cancer was demonstrated to have a gene expression profile that was highly similar to that of BRCA1 mutants, suggesting a possible defect elsewhere in the BRCA1 pathway.

In a similar vein, we recently identified a gene expression signature capable of discriminating p53 mutant and wild-type tumours with reasonable accuracy. Closer observation revealed that the p53 wild-type tumours with expression signatures more related to that of the mutants possessed significantly lower levels of the p53 transcript, as well as lower transcript levels of a number of p53 direct target genes (compared with the other wildtype tumours), suggesting the possibility of p53 deficiency in these otherwise p53-sequence-wild-type tumours. Moreover, Cox’s proportional hazards analysis confirmed a significantly poorer disease-specific survival in this subset of tumours compared with the other p53 wildtypes that was consistent with the survival curve seen in p53 mutants compared with all wild-types [32]. These findings highlight the intriguing view that mechanistically derived gene expression signatures may reflect the operational configuration of clinically relevant pathways in cancer, and thus may provide more robust prognostic information than the biomarkers from which they are derived.

The fundamental importance of receptors for breast cancer biology was studied in 58 node-negative breast cancers, where the expression of ERα markedly discriminated tumours into an ERα-positive subgroup and a corresponding negative group with associated genes [33]. The separation of breast cancers into ERα-positive and -negative subgroups using microarray profiling has also been demonstrated by a number of other
groups [21, 27, 29, 34]. ERα gene expression has recently been analysed together with ERβ, aiming at understanding their functional interactions in human breast cancer [35]. It was revealed that ERβ had an inhibitory effect on tumour cell proliferation and that the gene expression signature linked to low ERβ expression was associated with better disease outcome.

**Microarray profiling for therapy prediction**

Gene expression profiling has also been used to identify gene patterns that potentially identify genes related to response to certain cytotoxic agents like docetaxel [36]. This discriminatory array profile was based on a 92-gene set. In this limited series, one patient had progressive disease. The authors defined a positive tumour response as 25% or less residual disease, and resistant disease was defined as 26% or more residual disease. Conventional UICC and RECIST criteria define a 50% and 30% reduction, respectively, for a response [37]. If the UICC criteria had been applied, some of the tumours in the gene expression-predicted resistant group would have belonged to the UICC-responding group, despite the fact that they had a completely different expression profile for the selected gene signature.

Figure 1. The upper dendrogram shows status of TP53 in the five (six) subgroups identified in all 85 samples analysed by Sorlie et al. [28], indicated by the colour of the terminal dendrogram line. Red lines indicate tumours with mutated TP53 genes, green lines wild-type (WT) TP53 and black lines samples not tested. The lower dendrogram shows hierarchical clustering analysis using the 51 Norway carcinomas, three benign tumours and four normal breast tissues. It closely resembles the upper dendrogram. The subgroups are coloured accordingly and show that the group of tumours highlighted in orange changed position compared with the upper dendrogram. Furthermore, the basal-like tumours shown in red are inserted in between luminal subtypes A and C. To the left are shown the correlation coefficients for the dendrogram branches. Reproduced with permission.
Critical review of gene expression profiling

A Swedish group analysed clinical, histopathological and cell biological parameters in the previously discussed Amsterdam cohort (i.e. 97 patients) in order to determine whether these conventional variables were able, or not, to provide the same level of prognostic information as the microarray expression signature described by the authors [21, 42]. The Swedish group demonstrated a very similar outcome for both the good prognosis group and poor prognosis group using artificial neural networks, as well as the Nottingham prognostic index, compared with gene expression profiling. This suggests that empirically determined expression signatures may simply recapitulate the prognostic aspects of certain conventional factors, and that gene selection methods are needed to derive signatures with strong independent prognostic capabilities.

How do microarrays function?

The Affymetrix and Agilent systems are among the dominating microarray platforms on the commercial market (Affymetrix Inc., Santa Clara, CA, USA; Rosetta Inpharmatics, licensed to Agilent Technologies, Palo Alto, CA, USA). Affymetrix chips are synthesised by photolithography and solid-phase chemistry. Each perfect-match oligonucleotide is coupled with a mismatched oligonucleotide to measure non-specific hybridisation. Several (11–20) pairs of 25mer oligorobes are used for each gene or transcript. These bind to different regions of the cRNA, and improve signal-to-noise ratio and the dynamic range of detection. The multiple oligonucleotides also reduce cross-hybridisation effects. Biotin-labelled antisense cRNA amplified from sample mRNA is hybridised to the chip.

For the Agilent technology, antisense cRNA or cDNA from sample and reference are labelled with the fluorescent dyes Cy3 and Cy5 and hybridised simultaneously to the array (Figure 2). Disadvantages with the labelling method may be the two-colour dye bias and the sensitivity of Cy5 to decay with exposure to atmospheric ozone, although Agilent recently modified the wash protocol to reduce the problem with ambient ozone levels. The Agilent arrays are synthesised by inkjet technology. The array contains between 11,000 and 44,000 spots with millions of copies of a 60mer oligonucleotide or cDNA. Each gene or transcript is represented by only one 60mer. Agilent allows for flexible design, with possibility to choose between customer-designed probes between 25–60 bases long and 60mer probes designed by Agilent. The 60mer oligo format is advantageous relative to short 25mer length oligos in terms of sensitivity; however, shorter oligos are considered to be more specific. Agilent arrays are more tolerant of mismatches, which results in simplified analysis of polymorphic regions, but on the other hand decreases specificity.

Technical aspects of gene expression profiling: what can go wrong

RNA from breast cancer samples have been studied on different microarray platforms, utilising either long cDNA clones as probes [28], Affymetrix 25mer oligo probes [6] or 60mer oligo probes [21].

The diversity of microarray platforms utilised for gene expression analysis and the variability in microarray data emphasise the need for quality assurance. High quality of the RNA samples is essential for microarray analysis, and each RNA preparation must be rigorously assessed for quality. Careful measures need to be undertaken during all steps of the RNA extraction method to prevent the RNA from degrading and to maximise the yield. DNase treatment should be
included in the protocol to ensure the absence of contaminating DNA. The addition of proteinase K during RNA extraction markedly increases the yield [43]. Usually, RNA integrity is controlled for by measurement of the 28S/18S ribosomal RNA ratio on a capillary electrophoresis system such as the Agilent 2100 Bioanalyzer (Figure 3).

A number of factors influence the correlation between different microarray platforms and complicate cross-platform comparisons. Different factors may affect the variability in microarray gene expression results and analyses: technical factors such as microarray manufacturing, cDNA or oligo probe type, selection of genes represented on the arrays, sample collection, RNA extraction, synthesis of cDNA or cRNA, labelling with fluorescent dye, and hybridisation; instrumental parameters such as image acquisition and quantitation; computational factors including preprocessing, normalisation and analysis of data; and biological factors [44–46].

cDNA clones and oligonucleotides differ in their hybridisation properties. GC content, probe sequence length, degree of cross-hybridisation and average signal intensity influence the correlation between microarray platforms. The length of the cDNA clones makes them more prone to hybridise to homologous sequences on other genes, while short oligonucleotides are more specific. Alternatively spliced transcripts that differ in their hybridisation ability to short oligo- or long cDNA-based arrays may contribute to the differences between the platforms [46]. The Affymetrix platform is perhaps the most standardised and evaluated microarray platform, with high reproducibility; however, some studies have reported a lower reproducibility (using the same RNA preparations) [47–51]. Yauk et al. [47] reported good reproducibility for Agilent cDNA and 60mer oligonucleotide arrays. Poor correlation between Affymetrix chips and customised cDNA microarrays have been reported in several studies, while use of only sequence-verified and -matched genes for comparison of Affymetrix and Agilent

---

**Figure 2.** The principle of the Agilent microarray analysis. Reference and tumour RNAs are labelled with Cy3 and Cy5 fluorescent dyes, respectively, then pooled together and hybridised to the array. The relative abundance of mRNA from tumour versus reference is obtained from the Cy5/Cy3 intensity ratios. MMLV-RT, Moloney murine leukemia virus reverse transcriptase.
cDNA microarrays resulted in a better correlation [52–56]. Jarvinen et al. [46] observed a better correlation between Affymetrix and Agilent cDNA than with custom-made cDNA arrays. Better concordance has also been reported for Affymetrix and Agilent oligo or customised oligo microarrays [57, 58].

Conclusions
RNA expression profiling of human breast cancer can separate patients into subgroups with good prognosis signatures versus poor prognosis signatures. Different prognostic gene expression signatures with very little overlap have been reported in breast cancer, underlining the importance of improved analytical procedures and the use of large numbers of patients for validation. Gene signatures will likely be used in the future for optimised and tailored drug selections.

Acknowledgements
Our array studies have been supported by grants from the Swedish Cancer Society, the Cancer Society in Stockholm and the King Gustav V Jubilee Fund, Stockholm, Singapore Agency for Science and Technology and Research (A* STAR) and Bristol-Myers Squibb.

References


11. Smith I, Dowsett M, on behalf of the IMPACT, Royal Marsden Hospital, United Kingdom. In Comparison of anastrozole vs tamoxifen alone and in combination as neoadjuvant treatment of estrogen receptor-positive (ER+) operable breast cancer in postmenopausal women: the IMPACT trial. In Lippman ME (ed.). 82. 26th Annual San Antonio Breast Cancer Symposium 2003; S6, (Abstr 1).


