Microarrays in breast cancer research and clinical practice – the future lies ahead

Sofia K Gruvberger-Saal1, Heather E Cunliffe2, Kristen M Carr3 and Ingrid A Hedenfalk2,4

1Institute for Cancer Genetics, Columbia University, New York, New York 10032, USA
2Translational Genomics Research Institute, Phoenix, Arizona 85004, USA
3Joan and Sanford I. Weill Medical College, New York Presbyterian Hospital, Cornell University, New York, New York 10021, USA
4Division of Oncology, Department of Clinical Sciences, Lund University, SE-221 85 Lund, Sweden

(Requests for offprints should be addressed to I A Hedenfalk; Email: ingrid.hedenfalk@med.lu.se)

Abstract

Molecular profiling for classification and prognostic purposes has demonstrated that the genetic signatures of tumors contain information regarding biological properties as well as clinical behavior. This review highlights the progress that has been made in the field of gene expression profiling of human breast cancer. Breast cancer has become one of the most intensely studied human malignancies in the genomic era; several hundred papers over the last few years have investigated various clinical and biological aspects of human breast cancer using high-throughput molecular profiling techniques. Given the grossly heterogeneous nature of the disease and the lack of robust conventional markers for disease prediction, prognosis, and response to treatment, the notion that a transcriptional profile comprising multiple genes, rather than any single gene or other parameter, will be more predictive of tumor behavior is both appealing and reasonable. Promising results have emerged from these studies, correlating gene expression profiles with prognosis, recurrence, metastatic potential, therapeutic response, as well as biological and functional aspects of the disease. Clearly, the integration of genomic approaches into the clinic lies in the near future, but prospective studies based on larger patient cohorts representing the whole spectrum of breast cancer, oncogenic pathway-based studies, attendant care in bioinformatic analyses and validation studies are needed before the full promise of gene expression profiling can be realized in the clinical setting.

Endocrine-Related Cancer (2006) 13 1017–1031

Introduction

In recent years, microarrays have been used extensively to study molecular differences among different types of cancer. One of the most intensely studied tumor types in this context is breast cancer. Characteristic patterns of gene expression have emerged, reflecting molecular differences between previously known as well as newly defined subtypes of breast cancer. These molecular differences have been shown to correlate with clinical features, such as survival, prognosis, and treatment sensitivity, as well as traditional histopathological parameters. This review will focus on how the application of microarray-based technologies in the investigations of molecular markers has evolved from descriptive biological definitions of the tumor landscape to more functional and mechanistic studies with correlations to clinically important traits within the field of breast cancer management. Here, we summarize and evaluate the recent use of microarray technologies in their ability to define a molecular taxonomy of breast cancer and identify molecular predictors of disease progression and clinical behavior.

Microarray technology

Array-based molecular profiling represents a technological breakthrough in how biological samples, for instance tumor specimens, can be analyzed. Since the technology allows for the simultaneous analysis of many thousands of individual genes in cell or tissue samples, the complex biology of cancer may be studied much more comprehensively than previously possible. Present technology allows for microarray features to be genome representative (the human genome sequence
suggests that the total number of genes is approximately 20,000–25,000 (Consortium 2004). Hence, it is nowadays possible to monitor gene expression or copy number levels of almost all known human genes in a biological specimen in a single experiment. The impact of this technology defines a revolution in basic biological science, and once fully standardized and accurately validated is likely to revolutionize the practice of medicine.

While various microarray platforms exist, utilizing diverse manufacturing, labeling, and analysis methods, the general principle of a gene expression microarray experiment remains the same. Labeled cDNA or cRNA pools (conventionally termed ‘targets’) derived from samples of interest are allowed to hybridize to elements on the microarray (‘probes’ or ‘reporters’), each derived from the sequence of individual genes and located at a unique position on the array surface. The degree of fluorescence at each probe is quantitated, and represents the abundance of that specific gene transcript, enumerated as either a ratio to a reference sample or as an absolute intensity value. The utilization of high-density cDNA microarrays for profiling gene expression was first described in 1987 (Augenlicht et al. 1987). However, it was not until advances in library construction and sequencing and florescent-labeling technologies, that work from the laboratory of Dr Pat Brown at Stanford University in 1995 popularized the two-color robotically printed microarray on glass substrate method (Schena et al. 1995), which was soon after applied to cancer specimens (DeRisi et al. 1996). Expression profiles derived from printed cDNA and oligonucleotide (25–80-mer) microarrays, the latter of which can also be synthesized in situ (Hughes et al. 2001), are based on the comparison between a sample of interest and a common reference sample. Hence, results are ratio-based and the output depends on the reference sample used. Whereas initial studies commonly used a single cancer cell line, or even normal tissue from the cell type being studied, most contemporary investigations utilize a pooled reference representing either all samples included in the study or a mixture of cell lines. The pooled alternatives ensure better coverage in terms of genes expressed on the microarrays (i.e. greater information content from each experiment), with the mixture of cell lines perhaps being the most advantageous as it is more easily renewed to allow for accrual of new hybridizations to a study over time. However, the photolithographically in situ manufactured short (25 bp) oligonucleotide microarrays do not use a co-hybridized reference sample (Lockhart et al. 1996); hence, gene expression levels represent absolute values. There are advantages as well as disadvantages in using either cDNA, longer oligonucleotide, or short oligonucleotide microarrays but this is an issue beyond the scope of the present review, although recent cross-platform studies indicate that, if performed in an experienced laboratory, the three approaches generate reasonably concordant data (Irizarry et al. 2005, Larkin et al. 2005). cDNA (Pollack et al. 1999, Heiskanen et al. 2000) and more recently oligonucleotide microarrays (Lucito et al. 2003, Carvalho et al. 2004), microarrays of bacterial artificial chromosome (BAC) clones (Pinkel et al. 1998), and specifically designed oligonucleotide clones (Barrett et al. 2004; see also e.g. products from Agilent Technologies Inc., Santa Clara, CA, USA and Nimblegen Systems Inc., Madison, WI, USA) are now also being used to perform microarray-based comparative genomic hybridization (CGH) for high resolution DNA copy number analysis. Moreover, oligonucleotide single nucleotide polymorphisms (SNP) arrays are becoming more popular as they can report both copy number and genotype in the same hybridization (Bignell et al. 2004, Zhao et al. 2004). The recent introduction of application specific arrays (including exon-specific arrays and arrays for chromatin immunoprecipitation) together with technical advancements of established platforms presents a variety of options to conduct basic as well as translational research.

Data analysis

To date, a multitude of sophisticated methods and algorithms have been developed for the comprehensive analysis of microarray data. The approaches can coarsely be divided into supervised and unsupervised methods. Supervised analyses require prior knowledge, biological or clinical, to discriminate genes distributed between sample groups with statistical probability. Unsupervised methods are used to reveal the underlying variance structure of a dataset, typically utilizing pattern-recognition algorithms to define groups of samples with similar global molecular profiles. These methods thus minimize a priori assumptions about the data, and thus identify structures in array data without regard to known sample parameters. While unsupervised analyses are effective in classifying tumors that have similar molecular patterns for a large number of genes, such analyses are less effective at identifying altered profiles of small numbers of discriminatory genes that nonetheless correlate with important parameters, and are poorly suited for identifying prognostic variables (Simon et al. 2003). The identification of subtle differences that may help refine tumor classification, or find correlations with, for
example, response to a given therapy, dictates the need for supervised approaches. Critical measures of validity in array-based molecular classification include technical and/or biological replication and subsequent validation. Internal validation of a statistical model requires a training set and a separate test set that should not be used to develop the model to avoid overfitting of the data to the cohort under study. External validation, using an independent set of tumors is also an essential part of model development. Many review articles have been published recounting the various approaches and analytical methods used in microarray data analysis (e.g. Quackenbush 2001, Ringner et al. 2002, Simon et al. 2003, Ahmed & Brenton 2005, Allison et al. 2006), therefore they will not be discussed in detail here. In brief, there are various issues to consider; for example, in supervised classification analyses, data overfitting/non-generalizability is a typical problem in underpowered study designs with too few training samples and complex classification models. Cross-platform comparison of related datasets in silico is another ongoing challenge (Bammler et al. 2005); there has been some convergence in experimental protocols and analytical methods most commonly used, new algorithms to make disparate datasets more comparable (Shen et al. 2004), as well as data reporting guidelines and file format standards (Minimal Information about a microarray experiment (MIAME) and Microarray Gene Expression Markup Language (MAGE-ML)) (Brazma et al. 2001, Spellman et al. 2002), all of which will help overcome translation of genomic research discoveries for clinical utility. Moreover, sophisticated microarray databases, analysis tools, and pathway, and gene regulatory network software are now available both commercially as well as academically, enabling biological interpretation of complex genome-wide microarray data with greater efficiency (e.g. Gene Ontology, Ashburner et al. 2000; GenMAPP, Dahlquist et al. 2002; SMD, Sherlock et al. 2001; TM4, Saeed et al. 2003; BASE, Saal et al. 2002; GSEA, Subramanian et al. 2005; GenePattern, Reich et al. 2006; Ingenuity Systems, www.ingenuity.com; MetaCore, www.genego.com; and PathwayStudio, www.ariadnegenomics.com).

**Distinguishing tumors on the basis of their gene expression profiles**

Early microarray studies were focused on molecular definition of already known subtypes of cancer; for example, unique and different expression profiles exist between estrogen receptor α (ERα) positive and negative breast cancers (Perou et al. 2000, Gruvberger et al. 2001, West et al. 2001), and breast cancers derived from BRCA1 and BRCA2 mutation carriers (Hedenfalk et al. 2001). Whereas these findings may not have been unexpected *per se*, given the known phenotypic and biologically significant differences, these first studies illustrated the power and sensitivity of microarrays for molecular characterization. To date, gene expression profiles associated with, for example, ER protein expression, histological grade, lymph node status, ERBB2 (HER-2/neu) gene amplification, *p53* mutational status, inflammatory breast cancer, and carcinoma-derived stromal signatures have been defined (Ahr et al. 2001, Gruvberger et al. 2001, Sorlie et al. 2001, 2003, West et al. 2001, 2005, Iwao et al. 2002, van’t Veer et al. 2002, Dressman et al. 2003, Sotiriou et al. 2003, 2006, Bertucci et al. 2004, Bieche et al. 2004, Gruvberger-Saal et al. 2004, Miller et al. 2005). Identification of the most prominent genes constituting the expression profiles associated with already known subtypes has significantly improved our understanding of the complex biology involved in the development of these different tumor subtypes and facilitated identification of new potential therapeutic targets. Moreover, class comparison studies using gene expression profiling highlight the capacity of this technology to refine clinical diagnostic or prognostic tools, e.g. histological grade. These interesting approaches will be discussed further in sections below.

**Finding novel disease entities within breast cancer**

Breast cancer is morphologically and genetically highly heterogeneous, a belief that has been confirmed and strengthened in gene expression studies performed to date. Hence, traditional methods for subgrouping breast cancers relying on pathological and clinical data can only partially reflect the clinical diversity of the disease. Molecular profiling has been shown to be well suited to phenotypic characterization of breast cancer and potentially to discover new molecular classes among cancers with similar histological appearance. Several landmark microarray studies demonstrate that one can build a molecular taxonomy of breast tumors using this technology. The most evident stratification of breast tumors is across ER status, reflecting the vastly different biology inherent in these tumor subtypes (Perou et al. 2000, Gruvberger et al. 2001, West et al. 2001, van’t Veer et al. 2002). The ER status of a tumor clearly has a remarkable impact on the genes expressed by the tumor, and gene expression data can be used not only to classify ER positive versus ER negative tumors, but also to predict, on a
continuous scale, the actual level of the ER protein as determined by enzyme immunoassay (Gruvberger-Saal et al. 2004). Interestingly, when removing several hundreds of the key genes in the ER expression profile, it is still possible to predict the ER status of individual tumors, suggesting that the expression profiles associated with ER involve a large number of genes (Gruvberger et al. 2001, Gruvberger-Saal et al. 2004). Clinically relevant novel subgroups within the ER positive and ER negative breast cancers have also been identified in primary tumors (Perou et al. 2000, Sorlie et al. 2001). Initial studies suggested that ER negative tumors encompass three subgroups, one overexpressing ERBB2, one with tumors expressing genes characteristic of basal epithelial cells, and one with a gene expression profile similar to normal breast tissue. The ER positive group was characterized by higher expression of a panel of genes associated with luminal epithelial cells, suggesting that ER positive tumors commonly arise from this cell type. The different potential origins of the tumor cells (luminal versus basal epithelial cells) was supported by immunohistochemical staining for cytokeratins typical for these two cell types, keratins 5/6 and 17 for basal-like tumors and keratins 8/18 for luminal-like tumors (Perou et al. 2000). Tumors evolving from basal epithelial cells are thought to be a less common event, comprising 3–15% of all cases. Recent literature indicates that basal-like tumors, which include tumors null for BRCA1, are still a heterogeneous grouping requiring further subclassification, as they display a complex pattern of expression of basal-associated markers, including vimentin, p63, CD10 α-smooth muscle actin, and EGFR (Abd El-Rehim et al. 2004, Livasy et al. 2006). The subgrouping of ER positive/luminal epithelial tumors appears to further separate into two, or possibly three, subgroups, named A, B, and C respectively (Sorlie et al. 2001); however, the stability of these groupings is not clear. The luminal subgroup A demonstrates higher levels of expression of genes in the ERα gene cluster compared with tumors in subgroups B and C. On the other hand, the latter subgroups are characterized by high expression of a novel set of genes with unknown function but that seem to be shared by the basal-like and ERBB2 + subtypes. The existence of these molecularly distinct subgroups in breast cancer, especially in the ER negative subgroup, has been corroborated in other breast cancer gene expression datasets (Sorlie et al. 2003, Sotiriou et al. 2003). Moreover, it is likely that further clinical stratification will occur (the existence of an ‘apocrine’ breast cancer subtype with increased androgen signaling has recently been suggested; Farmer et al. 2005), with key advancements needed in predicting the activity of oncogenic pathways as well as identifying markers for response to specific conventional and molecularly targeted therapies. Continued efforts in integrating and validating data from related studies seamlessly with well-defined clinicopathological markers are required to accomplish a more refined, robust gene expression-guided method of breast tumor subclassification.

Microarrays have also been successfully used to identify distinct subgroups among familial breast cancers. Germline mutations in the BRCA1 and BRCA2 genes, which together account for a significant portion of hereditary breast cancers, have been shown to leave a characteristic imprint on the panel of genes expressed by the tumors (Hedenfalk et al. 2001). While BRCA1-derived breast cancers display certain histopathological characteristics that may aid in their characterization as BRCA1 tumors, they do not constitute an entirely uniform group. Moreover, BRCA2 breast cancers make up a considerably more heterogeneous group. Mutation screening techniques used in oncogenetic clinics are time consuming and expensive, thus the realization that tumors derived from individuals with BRCA1 or BRCA2 mutations each display characteristic gene expression profiles (Hedenfalk et al. 2001) opens the possibility for novel screening strategies (e.g. a BRCA1/2 diagnostic gene chip). In addition, extended knowledge of the defects causing the development of breast cancer may greatly improve both treatment strategies and intervention approaches for affected individuals in these high-risk families. The occurrence of a BRCA1-like gene expression profile due to methylation-mediated silencing and not a germline mutation also highlights the convergent utility of expression profiling and illustrates the high degree of sensitivity of this technology to identify defects in individual genes. Further analysis of these defects is likely to shed light on the functional relationship between specific genetic or epigenetic alterations and disease.

Microarray analysis of the histopathologically heterogeneous group of non-BRCA1/2 (BRCAx) familial breast cancers has also shown promising results in discovering novel subgroups (Hedenfalk et al. 2003). BRCAx tumor profiles appear to indicate the presence of multiple underlying alterations, which is most likely a reflection of their phenotypic heterogeneity. Not surprisingly, this heterogeneity has confounded linkage analysis studies to identify further cancer-predisposing mutations in BRCAx families. Although in many cases frozen tumor material is unavailable, recent studies have demonstrated the power of
expression profiling in this context and provide promising data to suggest that familial BRCAx tumors can indeed be subclassified into homogeneous subsets, separate from BRCA1 and BRCA2 tumors (Hedenfalk et al. 2003). Furthermore, microarray-based CGH from the same tumors revealed that these subgroups were each associated with specific somatic genetic alterations, supporting the hypothesis that there are multiple distinct subclasses of BRCAx tumors (Hedenfalk et al. 2003). Although the numbers of tumors in many of the above-mentioned studies are small, these initial studies illustrate that molecular profiling is a useful tool in the identification of distinct and homogeneous subclasses.

**Prediction of prognosis and response to therapy**

Molecular markers hold great promise for refining our ability to establish early diagnosis, prognosis, and prediction of response to therapy. Although microarray technology is providing significant optimism in this regard, many of the landmark reports have limitations and problems that are still being debated, based on experimental design, inappropriate statistical methods, or lack of validation (see reviews (Simon et al. 2003, Simon 2005)). Two seminal complementary class discovery studies first describing the use of array-based molecular characterization in breast cancer (introduced in the previous section) identified at least five specific subtypes, recognizable solely on differences in gene expression patterns (Perou et al. 2000, Sorlie et al. 2001). Importantly, these molecular subtypes correlated with clinical outcome (metastasis free as well as overall survival). Substantial differences in prognosis correlating with molecularly distinct subgroups highlights the importance of leveraging this information to work towards individualized therapy. In a report by Sorlie et al. for patients treated with adjuvant tamoxifen, it was suggested that this subgrouping is only useful for predicting prognosis following endocrine treatment. Importantly, however, when their profile was applied to a different dataset with most patients not receiving adjuvant treatment, distinct tumor subtypes still appeared to be associated with different survival although the correlation was less prominent (Sorlie et al. 2003). These studies, while promising, still provide misconceptions due to serious concerns over too few patients, heterogeneous groups of patients, tumors, and treatment regimens, and lack of external validation. These are the same problems that generally hamper biomarker trials, and microarray studies do not seem to be an exception (Ransohoff 2004, Simon 2005).

Where the above studies used an unsupervised correlative approach to identify clinically relevant breast cancer subtypes, supervised classification methods have been used to isolate sets of genes associated with defined clinical parameters such as outcome. This involves first ranking a set of genes defined as variant with respect to their strength of association to a clinical parameter, then building a mathematical model to predict the parameter using gene expression values as input. The ultimate usefulness of any prognostic predictor is to more accurately identify patients who need further treatment without over-treating those who are cured by surgery alone. Various prognostic questions have been addressed in breast cancer using gene expression microarrays (Sorlie et al. 2001, Ahr et al. 2002, van de Vijver et al. 2002, van’t Veer et al. 2002, Chang et al. 2003, Huang et al. 2003, Sotiriou et al. 2003, 2006, Wang et al. 2005); however, debate remains over the validity of the predictors defined in each of these studies. In the study by van’t Veer et al., poor prognosis, specified as occurrence of distant metastases within 5 years, was predicted with an accuracy of 83% in a group of 98 lymph node-negative patients under the age of 55 years, most of whom did not receive adjuvant systemic treatment (van’t Veer et al. 2002). The results were later confirmed in a larger set of tumors (295 tumors, 197 of which were not included in the previous study from which the predictor was generated) that were node-negative and -positive and who had or had not received adjuvant systemic therapy (van de Vijver et al. 2002). According to the authors, the 70-gene predictor from the combined data performed better than any single standard univariate clinically prognostic marker in predicting occurrence of distant metastases as well as overall 10-year survival (hazard ratio 4.6; 95% CI: 2.3–9.2 using Cox proportional hazards analysis). The prognosis predictor was associated with some of the traditional prognostic markers: the age of the patient, the histological grade of the tumor, and ER status, whereas there was no association to some other established prognostic markers such as number of positive lymph nodes, tumor size, and vascular invasion. It is interesting that no correlation between lymph node status and the poor prognosis profile could be seen (patients with node-positive and node-negative disease were evenly distributed between the two prognostic groups). However, one should keep in mind that the distinctive characteristics of the cohort studied could explain some of these relationships. Other confounding problems with this study include the fact that 61 of the 295 tumors in the validation study (van de Vijver et al. 2006, van’t Veer et al. 2002).
et al. 2002) were also included in the initial study. This incompletely validates the poor prognosis predictor, causing ‘information leak’ due to possible positive bias. In addition, these two studies were performed exclusively on younger patients (<53 years at diagnosis) with stage I and II breast cancer only. This obviously raises questions as to whether this specific set of 70 genes is the most optimal for predicting prognosis for patients belonging to other subgroups. Not withstanding that their predictor performed better than any single conventional prognostic marker or the St Gallen (Goldhirsch et al. 2001) and NIH (Eifel et al. 2001) consensus criteria, a study re-analyzing the patient data from the Dutch study showed that the accuracy of the Nottingham prognostic index (Blamey et al. 1979) in predicting prognosis was similar to the 70-gene predictor (Eden et al. 2004). Additionally, the same group generated an artificial neural network-based prognostic classifier using conventional prognostic markers that was found to give an accuracy in prognosis prediction close to the performance of the gene expression predictor proposed by van’t Veer and co-workers (Eden et al. 2004), suggesting that the use of established conventional prognostic markers could be improved. However, expression levels of the 70 genes do appear to have prognostic value when measured using quantitative reverse-transcription PCR (QRT-PCR; Espinosa et al. 2005).

In a separate but related study by Wang et al., a 76-gene predictor was generated to predict development of distant tumor recurrence in 286 patients not receiving systemic treatment, with lymph node-negative primary breast cancer of variable size and from all age groups (Wang et al. 2005). This 76-gene predictor was determined to be the only significant variable in multivariate analysis for distant metastasis-free survival, again reported to outperform established prognostic factors. This predictor showed 93% sensitivity and 48% specificity in an entirely independent validation set of 171 lymph node negative patients with a hazard ratio of 5.55 (95% CI: 2.46–12.5) for distant metastasis in 5 years after adjustment for conventional prognosticators. Interestingly, comparison of the 76-gene predictor (Wang et al. 2005) and 70-gene predictor (van’t Veer et al. 2002) revealed only a three-gene overlap. Wang et al. suggested that this is likely to reflect differences in patient/sample characteristics and different array platforms used (Affymetrix versus Rosetta). Bioinformatic analyses of these datasets reveal confusion with respect to building prognostic models; for example, Ein-Dor et al. demonstrated the ability to compile several different prognostic signatures with similar performance from the NKI-Rosetta dataset (Ein-Dor et al. 2005). Shen et al. also reported a meta-analysis indicating that prognostic signatures derived from one cohort performed very differently in other cohorts (Shen et al. 2004). Given the plethora of commentary on these studies in the literature, the feasibility of using a reported prognostic predictor as a clinical test or for randomization into clinical trials remains a concern. Further efforts are needed to compile and jointly analyze related microarray datasets to potentially generate a consensus gene expression prognosis predictor from a larger patient population instead of having many different predictors based on smaller datasets. For example, if a prognostic signature was generated from the so-called validation dataset of 295 consecutive tumors from the Dutch study, would this be represented by the same top genes as in the 70-gene profile from their first study? In addition, why have class prediction studies to a great extent ignored the fact that one can first stratify tumors into molecular and or histopathological subclasses with more predictable clinical outcome? Even given the known genetic complexity of breast tumors, ER status, and tumor grade, perhaps more than any other parameters, appear to define genomic changes observed in breast cancer progression, reflecting distinct biology. It is therefore reasonable to make the assumption that predictive signatures would be more accurate and powerful if these known clinically prognostic stratifications were first taken into account (Simpson et al. 2005, Sotiriou et al. 2006).

Sotiriou et al. recently took a different approach when they built a predictive model for breast tumors of high versus low risk of recurrence among patients whose tumors were histological grade 2 (Sotiriou et al. 2006). Grade 2 tumors comprise 30–60% of all breast cancers, and as they are associated with an intermediate risk of recurrence, clinical decision making is difficult. The results suggest that the incorporation of molecular profiles into the histological classification of breast cancer may improve the accuracy of grading and thus its prognostic significance.

Although it is agreed that these studies have contributed to our overall understanding of breast cancer etiology and biology, it is important to note that a common problem of many first-generation prognostic predictors is that while the identified gene signatures may show some prognostic power, they give limited insight into the oncogenic pathways, which drive tumors to progress and metastasize. Some steps to break this ‘black box’ have been taken (Bild et al. 2006), but much more research in this area is needed before microarray-based research discoveries will fulfill their promise in the clinic.
A number of reports have now focused on the need to identify patients not only with a high risk of recurrence but also their predicted response to a given therapy. These studies have sought to identify gene expression profiles associated with tumors sensitive versus resistant to chemotherapeutic regimens, such as e.g. docetaxel (Chang et al. 2003, 2005, Iwao-Koizumi et al. 2005), adriamycin/cyclophosphamide (Sotiriou et al. 2002b), paclitaxel/fluorouracil/doxorubicin/cyclophosphamide (Ayers et al. 2004), and epirubicin/cyclophosphamide/paclitaxel (Modlich et al. 2004) in the neoadjuvant setting. These studies analyzed core biopsies (Chang et al. 2003, Modlich et al. 2004) or fine needle aspirates (FNAs; Sotiriou et al. 2002b, Ayers et al. 2004) from tumors before onset of treatment, and/or core biopsies, FNAs or surgically removed tumors after completion of neoadjuvant treatment. All these studies show promising results in identifying panels of genes suggested to be associated with response to treatment. Of interest, an underlying feature has emerged that patients with more aggressive tumors appear to respond more favorably to chemotherapy. In particular, Rouzier et al. have recently shown that the rate of pathological complete response (CR) in a neoadjuvant setting correlates with the intrinsic molecular subtypes previously described (Rouzier et al. 2005). Presumably due to the high correlation between molecular subtype and ER status, these subtypes were not independently associated with CR, as the microarray-based signature provided limited additional information in a multivariate analysis (Rouzier et al. 2005). Whether genes selected in these studies will turn out to be useful not only in the neoadjuvant setting but also for predicting long-term therapeutic response is still not known. Neither do we know whether these gene signatures could be informative for patients in other cohorts and for patients at different stages of the disease. However, the idea of being able to determine the optimal treatment for a patient from an FNA sample of a tumor taken before surgery is obviously very attractive. Several additional class prediction studies have addressed whether prognostic profiles correlate with a specific adjuvant treatment (Bertucci et al. 2002, Ma et al. 2004, Paik et al. 2004, Jansen et al. 2005). The latter studies report successful predictors for prognosis after treatment (progression-free survival) in patients with tamoxifen-resistant breast carcinomas. Again, few genes overlapped in related studies, highlighting the likely impact of unresolved confounding issues such as differing sets of genes included in the studies, patient cohorts having different characteristics, and different analytical methodologies used. It appears that not only the identity of genes used for predictors is a crucial factor, but their performance is also impacted by the way in which they are used in the analysis. Perhaps, more focus should be placed on how the different genes are weighted in the analysis, possibly even including conditional dependence between different sets of genes in the algorithm (Paik et al. 2004) in order to generate the most optimal predictors. Our current potential to accurately select the most optimal treatment regime for individual patients based on transcriptional profiles therefore remains unclear.

Clinical application of genomic research in breast cancer oncology

The use of gene expression profiling in breast cancer to predict clinical outcome and therapeutic response has yet to reach the stage where it can be optimally implemented in the clinical setting. A number of confounding issues remain, including differences in patient selection, array platforms and chemistry, and lack of consensus in use of analytical methodologies. For microarrays to become a generalized clinical tool, more efforts should be made to allow easier comparison between studies, by developing tools to map genes between platforms, and standardizing the data inter-change formats and methods. Ideally, to create an optimal classifier or predictor, the arrays from which the technique is developed should carry probes representing the entire human genome. Furthermore, the selection of the most optimal gene set to be included in a prognostic predictor should be performed using large sets of tumors from patients with the same treatment regimes, clinical stage, and subgroup, and perhaps the expression profiles should not consist of individual genes but rather of aggregate measures from groups of genes together, i.e. meta-genes representing specific contexts or activity states of known gene groups (Huang et al. 2003, Nevins et al. 2003). The 70-gene predictor identified by van’t Veer et al. performed well in their study (van’t Veer et al. 2002); however results from a multi-center validation study presented at the 2004 San Antonio Breast Cancer Symposium fell short of expectations. The Breast Cancer International Group (BIG) has launched an European Organisation for Research and Treatment of Cancer (EORTC)-led clinical trial called MINDACT (Microarray In Node negative Disease may Avoid ChemoTherapy). This trial is based on upfront patient stratification using the 70-gene predictor, now available as a commercial laboratory test called Mammaprint, developed and marketed by Agendia. A major obstacle to the translational aspects of the associated study is the necessity to use fresh frozen...
specimens, the proportion of tumor cells versus normal cells in the biopsy, and the requirement for expertise necessary to conduct sample hybridization to the prognostic array, which precludes the utility of this test in a routine clinical setting. Nevertheless, the results from this trial, recruiting 6000 patients, may help reconcile many existing questions and may reveal an improved prognostic gene set from analysis of a large cohort of patient samples.

In answer to important technical obstacles, research is now focused on ways to implement expression profiling from formal fixed paraffin embedded (FFPE) specimens. Arcturus recently released the Paradise Reagent System to enable extraction and amplification of RNA from FFPE samples, which can be used in conjunction with the Affymetrix Gene Chip XP3 arrays, enabling whole-genome expression profiling from archival material. Limitations in the robustness of this technology, such as the high degree of RNA fragmentation (to approximately 100 bp) have, however, restricted widespread application to date. While technical advancements are being made in this field, platforms that have the potential to profile archival samples that carry a wealth of clinical annotation will no doubt shorten the time for biomarker development (Abramovitz & Leyland-Jones 2006). In answer to this need, an alternative genome-based methodology that is suitable for implementation in clinical oncology is QRT-PCR from FFPE material. The Oncotype DX assay developed by researchers at Genomic Health is one example of how QRT-PCR is well suited as a clinical assay platform derived from microarray discovery studies. Oncotype DX distinguishes good from bad prognosis following adjuvant tamoxifen for patients, using a proprietary analysis of the expression of 21 known genes. This assay includes five reference genes, which show minimal variation among different breast tumors, and compensates for variation in absolute PCR signal resulting from variability in specimen quality. This test has been successfully applied to a retrospective cohort of 668 tamoxifen-treated, hormone receptor positive breast cancer patients (Paik et al. 2004). It employs a mathematical algorithm called the recurrence score (range 1–100) to calculate continuous risk for relapse and death for patients receiving adjuvant tamoxifen, and has recently been shown in a large population-based case-control study to be an effective predictive test for ER positive, node-negative breast cancer patients treated or untreated with tamoxifen and no chemotherapy (Habel et al. 2006). It will be of interest to evaluate the relative utility and cost effectiveness of the Oncotype DX assay compared to existing clinical tools such as Adjuvant! Online. Investigation is underway to evaluate the utility of the Oncotype DX assay to predict benefit from receiving chemotherapy (NSABP trial B-20). Results from this study presented at the 2005 American Society of Clinical Oncology meeting revealed a linear relationship between the risk of recurrence score and benefit from receiving chemotherapy, and thus a useful clinical marker (although the question remains regarding risk threshold for potentially forgoing chemotherapy). The trial assigning individualized options for treatment (TAI-LORx) trial has been launched to validate the usefulness of Oncotype DX as a predictor of chemotherapeutic response for patients with ER positive, node-negative tumors.

Questions remain for both the Mammaprint and Oncotype DX tests, whether the suggested gene expression-based predictors generated to date have been investigated in sufficient detail for applicability to diverse patient cohorts, from different institutions, and in comparison with conventional markers (Eden et al. 2004; www.adjuvantonline.com). The looming question remains whether costly clinical trials today may prevent the investment in and rigorous clinical validation of newer and more promising tests in the future. At this point, it seems reasonable to believe that molecular profiling could indeed be improved, and that confounding issues need to be addressed before steps are taken into clinical practice.

**Understanding the biology of breast cancer**

Not only can gene expression data be used to distinguish tumors into subtypes and facilitate prognostication, but is also particularly well suited to measure changes in gene expression patterns on a continuous scale (Gruvberger-Saal et al. 2004). Although the magnitude of gene expression changes observed in microarray analyses is blunted relative to actual levels, the underlying biological behavior of a tumor reflected in its gene expression signature is still a powerful illustration of the sensitivity and robustness of the technique to define pivotal oncogenic pathways such as those related to metastasis and development of drug resistance.

Metastasis, the spread of cancer cells from the primary tumor to remote sites of the body, is the main cause of death in cancer patients. The biological process of metastasis is still not fully understood, but the prevailing model has been that the ability to metastasize is acquired late in tumor progression in just a rare number of cells within the tumor. If this holds
true, the primary tumor should not necessarily display a specific metastasis-specific phenotype similar to the phenotype of its metastases. From a biological standpoint, the above-mentioned studies generating prognosis predictors from gene expression signatures of primary tumors challenge the previously accepted notion that metastatic potential is acquired during the multistep process of malignant progression. These studies now indicate that a patient’s metastatic risk is inherent in the gene expression pattern derived from the primary tumor. This is further supported in a recent study showing not only that prognostic signatures are maintained throughout the metastatic process, but importantly, so are their molecular subtypes (Weigelt et al. 2005). Further studies suggest that the majority of paired primary tumors and metastases, in the form of lymph node metastases or distant metastases from the same individual, are more similar with regards to gene expression signatures than tumors or metastases from separate individuals (Perou et al. 2000, Weigelt et al. 2003, Wang et al. 2005). These findings have led many to suggest that a primary tumor gains the genetic changes favoring metastasis early in development and that this ability to metastasize involves more than a few rare cells since it can be detected by studying the phenotype of the whole tumor. However, it remains possible that at least some of the changes that potentiate metastasis may be early changes that also promote non-invasive growth. Furthermore, given the large degree of heterogeneity among breast tumors, it is perhaps not unexpected that the phenotype of a primary tumor is retained in its metastases to the extent that it makes them more related to each other than to primary tumors from other individuals.

Identification of genes that may predict site of distant metastasis is also a subject under investigation. To this end, Kang et al. have suggested that there may be unique sets of genes predictive for this mechanism, but no evidence has been found that the tissue location of metastases influences the general gene expression profile of the primary tumor in humans (Kang et al. 2003). However, this could be due to limited sample size or it may support the hypothesis that metastatic cells are directed to specific locations of the body by bone marrow derived cells creating a pre-metastatic niche (Kaplan et al. 2005). In a microarray study of unmatched primary tumors and metastases from adenocarcinomas of varying tissue origins, a shared 17-gene predictor associated with metastatic potential was isolated suggesting that even though primary tumors and metastases from the same individual appear to be very similar, a common set of metastasis-associated genes can be identified within primary tumors prone to metastasize (Ramaswamy et al. 2003). However, these results are only the beginning of expression profiling in metastasis-related research and still include only a very small number of tumors.

Several class comparison studies have attempted to correlate gene expression signatures with biologically aggressive tumor phenotypes and with clinical stages of cancer progression. In a study by Ma et al. cells representing various stages of breast tumor progression, from normal ductal epithelium through carcinoma in situ to invasive carcinoma, were laser capture microdissected and subjected to gene expression profiling (Ma et al. 2003). The largest transcriptional variation was detected at the transition from normal to precancerous cells, suggesting that invasive potential may be present and detectable in premalignant lesions. While no molecular signatures correlating to the different stages of malignant growth could be determined, a correlation between distinctive gene expression profiles and histological grade was observed, suggesting that molecular profiles may be useful for improved grading of breast tumors (Ma et al. 2003), a conclusion also drawn by other groups (Sotiriou et al. 2006) as previously discussed. On the whole, the community appears to agree that heterogeneity in gene expression profiles reflects the large degree of biological and pathological heterogeneity among human breast tumors, which is certainly impacted by, e.g. hormonal and microenvironmental influences. The consequential and understandably profound alterations in gene expression may nevertheless be clarified using alternative genomic approaches. It is interesting to speculate whether a more consistent and robust means of identifying global molecular alterations associated with the different stages of breast cancer progression might be to analyze global genomic profiles using an array-based CGH approach.

Reports of the use of microarrays as a molecular readout from animal or cell line models are vast and have contributed significantly to our understanding of the biological behavior of breast cancer, underlying mechanisms of tumor progression, and development of drug resistance. Minn et al. used a mouse model to identify a gene expression signature associated with increased breast cancer metastasis to the lung, suggesting that at least in some cases, the site for metastasis is reflected in the genetic signature of the primary tumor (Minn et al. 2005). The combination of information from studies of primary breast cancers with in vitro cell line models may also provide insight into various aspects of tumor development and response to therapy. Culilffe et al. took this approach...
when they performed a comparative analysis of breast cancer cell lines treated with various regulators of growth and differentiation in vitro, and previously published tumor expression profiles (Cunliffe et al. 2003). This study highlighted important functional correlates between the response elicited by individual pharmacological agents and ER status or disease outcome in patient samples. Dai et al. similarly investigated the utility of a cell proliferation signature as a predictor of extremely poor outcome in a subpopulation of breast cancer patients (Dai et al. 2005). Recently, Oh et al. developed a gene expression-based outcome predictor for survival in hormone receptor positive breast cancers by correlating in vitro pharmacological expression profiles with biological tumor differences (Oh et al. 2006). Research of this nature demonstrates the power of linking profiling data from clinically annotated primary human tumors and in vitro model systems, together with functional, biological, and pharmacological data, and illustrates how the behavior of breast cancer cells in culture or in model systems can be linked to clinically relevant properties of human breast tumors.

Tumor heterogeneity and the presence of non-malignant cell types

In order to understand the process of tumor progression and to fully determine the characteristics of breast cancer, it is important to identify genomic profiles associated with the different cell subtypes that together make up the heterogeneous tumors or the location of neoplastic cells within a tumor mass. Molecular profiling of patient material enriched for tumor cells via FNAs (Sotiriou et al. 2002a, Pusztai et al. 2003), core needle biopsies (Ellis et al. 2002, Chang et al. 2003), laser capture microdissection (Zhu et al. 2003), and FFPE tumor sections (Paik et al. 2004) exemplify the use of new techniques to refine the accuracy of tumor-specific profiles. In an elegant study by Allinen et al. all the constituent cell types of distinct stages of breast cancer progression were isolated and purified, and gene expression profiles representing each cell type during progression were generated using serial analysis of gene expression (SAGE) (Allinen et al. 2004). Extensive changes, including overexpression of chemokines, were detected in all cell types suggesting that the interaction between constituent cell types is likely to play an important role in breast tumorigenesis. Another study tested the hypothesis that tumors resemble wounds and identified a so-called ‘wound healing’ signature, developed by treating fibroblast cell lines with serum and capturing a core set of responsive genes (Chang et al. 2004). The set of ‘wound healing’ genes was highly predictive of poor prognosis in multiple cancer types, including breast cancer, and the authors concluded that this signature identifies a biological feature of the tumor and its environment responding to the malignancy. In another effort, the same group analyzed two types of fibroblastic tumors to identify different populations or activation states of fibroblasts, and the stromal signature was then applied to discern breast cancers with significant differences in clinical outcome, suggesting that the stromal response in carcinomas is involved in tumor progression. While it is debatable whether signatures can be well ascribed to specific cell types when analyzing heterogeneous tumor samples including multiple cell types, these studies together provide important clues into the known role of the cellular microenvironment in tumorigenesis, and may also provide a resource for developing new molecular targets in non-epithelial, stromal cells for cancer intervention and treatment. Of note, whether the ‘wound healing’ signature corresponds to the tumor–host interaction or is simply predominantly a profile of the response to mitogenic growth factors in the serum is now questionable, particularly, after it was recently described that a combinatorial expression of MYC and CSN5 is a strong inducer of the ‘wound healing’ signature (Adler et al. 2006).

Future prospects

The use of gene expression profiling to address clinical issues clearly illustrates that the molecular signatures of tumors contain information regarding clinical behavior. However, the prognostic studies performed to date are extremely small in the context of evaluation of prognostic indicators and have only been applied to a subset of breast cancer patients with less advanced disease. Clearly, prospective studies based on larger patient cohorts representing the whole spectrum of breast cancer are needed before the full power of gene expression profiling will be realized in clinical medicine. Nevertheless, we are encouraged by results from studies using microarrays for classification, prognosis, and treatment prediction of breast cancer indicating that microarrays have the potential to be of direct benefit to the patient in the near future. These initial studies have been promising enough to provide a novel, personalized dimension to the treatment and care of breast cancer patients; however, development of predictive signatures within the dimension of molecular subtypes is an area requiring further investigation. It remains in debate whether studies to date have successfully established the utility of gene expression profiling to identify a larger group of patients with good prognosis, who may not benefit from adjuvant
systemic chemotherapy. Further studies aimed at elucidating the effect of different treatment regimens on disease outcome and/or molecular subtype, combined with efforts to develop targeted therapies, are needed to identify those patients most likely to benefit from available and novel adjuvant or neoadjuvant treatment regimens. In addition, microarrays provide a tremendous opportunity to explore the complex biology of breast cancer in detail and to potentially identify new targets for therapy. We must also be mindful of the fact that transcriptional events only partially correlate with protein levels, such that post-translational events may need to be taken into account, and also that certain events associated with malignant phenotypes are reflected on the DNA and protein levels. To conclude, researchers need to employ an array of different technologies and utilize the whole spectrum of biological and clinical data available in order to elucidate and clarify the multifaceted complexity of breast cancer. Translation of all forms of -omic data for clinical application is certainly a current research focus and remains a significant challenge. Rapid evolution of high throughput techniques in addition to consolidation of datasets and analytical methods will undoubtedly revolutionize and complement current patient management procedures. In an effort to assist patients with poor prognosis, the Molecular Profiling Institute (MPI, www.molecularprofiling.com) in consultation with the Translational Genomics Research Institute (TGen, www.tgen.org), launched Target Now in 2003. It is a molecular profiling tool designed for cancer patients for whom all standard therapies have failed, and aims to assess whether a drug target can be identified in an individual patient’s tumor. Initiatives of this nature highlight the urgency within the research community to maximize information gained from research discoveries for immediate clinical utility and patient benefit. Quintessentially, microarrays still hold the promise to improve rather than replace current prognostic and treatment predictive tools.

Acknowledgements

The authors would like to thank Professor Mårten Fernö for critical review of the manuscript.

Funding

This work was partly supported by the Swedish Cancer Society (IAH). The authors declare that they do not have any conflict of interest.

References


Sotiriou C, Khanna C, Jazaeri AA, Petersen D & Liu ET 2002a Core biopsies can be used to distinguish differences in expression profiling by cDNA microarrays. *Journal of Molecular Diagnostics* **4** 30–36.


