When should I start using a new biomarker: Focus on expression arrays?

C. Desmedt, C. Sotiriou*

Institut Jules Bordet, Functional Genomics and Translational Research Unit, Fonds National de la Recherche Scientifique (F.N.R.S.), Université Libre de Bruxelles (U.L.B.), Brussels, Belgium

Introduction

A tumour marker is a tool which enables the clinician to answer clinically relevant questions regarding a cancer disease. By definition, a marker represents a qualitative or quantitative alteration or deviation from the norm of a molecule (DNA, mRNA, protein), substance or process that can be detected by some type of assay [1]. The recent advances in basic research and genomics have improved our understanding of the biological basis of tumour development and progression. As a result, a variety of molecular tumour markers characterised in the lab, have been studied in the clinic for their potential to predict disease outcome or response to therapy. However, very few markers appear to provide definitive prognostic or predictive information. The majority of these clinical studies were retrospective and the results, so far, have not always been uniform with significant discordance between detection methods of marker expression.

DNA microarray-based technology has provided researchers with an ideal opportunity to begin taking steps towards performing comprehensive molecular and genetic profiling of cancer. The simultaneous study of thousands of genes, rather than focusing on just a few with traditional methods, has transformed the microarray technique into a powerful, holistic analytical tool. Clinical investigators rapidly saw the great potential of this technology, not only for gaining new insights into cell biology, but also as a powerful prognostic and predictive tool.

In this article, we will use breast cancer as an example to describe the state of the art of gene expression studies in identifying prognostic and predictive markers, their current limitations and future promises, and how molecular signatures might impact on individual patient management.

Gene expression signatures as a prognostic and predictive tool

Risk stratification for breast cancer patients based on currently used clinical guidelines is far from perfect and much progress is needed in identifying those patients who would really need adjuvant systemic therapy, and in identifying the best therapy for the individual patient. Thereby patients with similar clinical and pathological features have been reported with distinct clinical outcome and treatment response, suggesting that there exist additional molecular mechanisms determining tumour behaviour and progression. In this context, different research groups conducted comprehensive analyses of gene expression profiling data resulting in the identification of several gene expression signatures predicting clinical outcome [2–10] (Table 1). Interestingly, a common feature to all these gene expression signatures is that when they are compared to conventional clinico-pathological criteria, they all identified groups of low-risk patients not necessarily needing additional systemic adjuvant treatment. This implies that the number of patients that would receive unnecessary treatment could be significantly reduced, emphasising the advantage of these gene signatures over the clinical guidelines.

Although the prognostic potential of these different signatures is tremendous, it might be surprising that there is only little or no overlap between the different gene lists and an important question which then pops up is 'Do these individual signatures agree in predicting prognosis?' In an attempt to answer that question, Fan and colleagues compared five different expression profiles developed during the last 6 years [14]. Despite the differences in gene lists, in populations and platforms used, the majority of these classifiers seemed to predict accurately for independent data, highlighting an important observation that all these gene cassettes contain similar information with regards to the outcome of interest. Unfortunately, this study was only carried out on one dataset and did not investigate the biological foundation of these concordant results. To further investigate this question, our group in collaboration with the Swiss Institute of Bioinformatics, conducted very recently a meta-analysis of publicly available gene expression data of nearly 3000 breast cancer patients and showed that: (1) most prognostic signatures have similar prognostic power, which is largely attributable to proliferation,

Table 1
Main gene expression profiling studies with prognostic implications in breast cancer

Signature	Type of microarray platform	Samples (Characteristics)	Main findings	Ref.
'Fishing-expedition' approach				
70-gene prognostic signature	Long-oligonucleotides (Agilent Hu25K – 24,479 oligonucleotides)	78 (all <55 yrs old, LN– and untreated)	Prediction of clinical outcome, defined as presence of distant metastases at the 5-year mark	[2]
		295 (including 61 from previous study – both treated and untreated – 151 LN– and 144 LN+)	Validation of [2]	[3]
	Long-oligonucleotides (Custom-designed 'Mammaprint®')	307 (T1–T2 tumours from pts <61 yrs old, LN–, and untreated)	Independent validation of [2] and [3]	[11]
76-gene prognostic signature	Short-oligonucleotides (Affymetrix U133A – 22,283 probe-sets)	286 (untreated LN– pts – training set 115; validation set 171)	Prediction of clinical outcome, defined as presence of distant metastases at the 5-year mark	[4]
(60 genes for ER+ and	Short-oligonucleotides (Custom-made Affymetrix VDX2 – 297 genes)	180 (untreated LN–)	Validation of [4]	[5]
16 genes for ER– tumours)	Short-oligonucleotides (Affymetrix U133A – 22,283 probe-sets)	198 (T1–T2 tumours from pts <61 years old, LN–, and untreated from ref. [12])	Independent validation of [4] and [5]	[13]
'Hypothesis-driven' approach				
Fibroblast Core Serum Response list (512 genes)	cDNA (43,000 features)	Fibroblasts from ten anatomic sites	Identification of similarities between tumours and wounds	[6]
	Long-oligonucleotides (Agilent Hu25K – 24,479 oligonucleotides)	295 (from ref. [14])	Robustness, scalability and integration of this signature to predict survival	[7]
Gene Expression Grade Index (97 genes)	Short-oligonucleotides (Affymetrix U133A – 22,283 probe-sets)	189 (training set: 64 ER+ pts; validation set: 125 untreated pts) + 3 publicly available datasets (from refs [2,4,14])	Identification of Genomic Grade associated with histological grade able to classify histological grade 2 pts into high and low risks of recurrence	[10]
	Short-oligonucleotides (Affymetrix U133A – 22,283 probe-sets)	335 (ER+ both untreated and Tam-treated – 86 from ref. [15]) + 4 publicly available datasets (from refs [2,4,14,16])	Definition of distinct subtypes of ER+ tumours with the Genomic Grade	[17]
Death-from-cancer signature (11 genes)	Short-oligonucleotides (Affymetrix U95Av2 – 12,625 probe-sets)	Mouse/human comparative approach for the identification of the signature and validation on 1153 cancer patients of whom 97 were breast cancer patients	Identification of an 11-gene stem-cell resembling signature	[8]
Invasiveness Gene Signature (186 genes)	Short-oligonucleotides (Affymetrix U133A and B – ~45,000 probe-sets)	Comparison of six CD44+/CD24– or low cells isolated from breast cancers with normal epithelial cells from three mammoplasties + Validation on ref. [3] and [4]	Identification of a tumorigenic breast cancer signature	[9]

CT: chemotherapy; LN: lymph node status; pts: patients; IFN: interferon; ER: oestrogen receptor.

and (2) that the prognostic power of these signatures is mainly limited to the ER-positive patients, as the majority of the ER-negative breast cancer are highly proliferative [16]. These results support the superiority of having a quantitative measurement of proliferation compared to the one obtained using semi-quantitative assays such as immunohistochemistry.

Although increasing breast cancer prognosis is of critical importance to better identify the patients not needing treatment, it is not sufficient since we also need to know which therapy will benefit the individual patient. Indeed, only a proportion of patients will respond to a particular treatment whereas most will experience the adverse side effects. Currently, the only accepted and recommended predictive markers for breast cancer are the hormonal receptors, which are used to select patients likely to respond to hormone therapy, and HER2 for selecting patients for treatment with trastuzumab.

Several predictors have been developed for endocrine therapy. Jansen and colleagues identified 44 genes that predicted for response to tamoxifen in the advanced disease setting [18], whereas the two-gene expression ratio, HOXB13 versus IL17BR and the Recurrence Score (RS) were identified to predict clinical outcome in primary breast cancer patients treated with adjuvant tamoxifen [12,19]. It is difficult, however, to dissect endocrine sensitivity from inherent prognosis. For example, it has been shown that the genomic grade index, (GGI), the RS and the two-gene expression ratio HOXB13/IL17BR were associated with clinical outcome in both systemically untreated and tamoxifen-treated populations [14,17,20,21]. This suggests that adjuvant tamoxifen monotherapy does not alter the poor clinical outcome of the high grade/high recurrence score/high two-gene ratio subgroup, respectively, which are most probably characterised by highly proliferating tumours. In the context of chemotherapy, few studies have been reported so far. The major reason is that most of these studies ideally require prospective sample collection in the context of a clinical trial. Several groups have identified genes associated with response to chemotherapy [22–27]. However, the statistical confidence of these studies remains low due to the small sample sizes and the unselected patient populations. Indeed, it has been shown in the literature that ER-negative and high grade tumours were associated with a better response to chemotherapy. It is therefore not surprising that a number of genes reported to be predictive of response are in fact associated with ER, grade or proliferation genes. Recently, using a 'hypothesis-driven' approach, Duke's researchers identified several expression patterns associated with the deregulation of a variety of oncogenic pathways that could predict response to different therapeutic agents targeting specific deregulated pathways [15]. The same group, using publicly available drug sensitivity data derived from *in vitro* experiments, developed multiple classifiers of response to a variety of chemotherapy drugs and showed that a combination of these classifiers accurately predicted response to preoperative multi-drug regimen treatments derived from two breast cancer studies [28].

Current limitations for transferring gene expression signatures from bench to bedside

Despite these interesting and positive studies, several doubts have been raised regarding the reliability of this new tool in clinical applications such as prognostication and treatment prediction. Indeed, this skepticism finds its roots in a number of microarray studies that investigated the same clinical problem (for example, the prediction of good outcome versus poor outcome) but generated different gene expression classifiers with only a small number of overlapping genes. There are several expected technical, analytical and biological reasons for these seemingly discrepant results.

Technology-related considerations

From a technical point of view, the assessment of an ideal prognostic or predictive test should be fast and robust with little inter-operator and inter-institution variability. One of the first formal investigations of this question compared expression data obtained from the same sample on three different microarray platforms [29]. Correlations in gene expression levels and comparisons for significant gene expression changes were calculated, and showed considerable divergence across the different platforms. Later, several research groups conducted similar comparitive studies and showed that concordance estimates improved dramatically after filtering for gene nucleotide sequence identity [30–33]. However, one disadvantage of this approach is that it only considers genes which map close or identically with probes used across platforms and therefore reduces drastically the gene expression information which is retained in the analysis.

In view of these concerns raised on one hand and the great potential of this technology for tailored medicine on the other, the U.S. Food and Drug Administration launched the Microarray Quality Control (MAQC[1]) project involving 137 participants from 51 academic and industry partners to systematically address the technical reproducibility of microarray measurements within and between laboratories, as well as across

[1] http://www.fda.gov/nctr/science/centers/toxicoinformatics/maqc/

different microarray platforms [34]. To this end, they generated gene expression data from four titration pools from two distinct reference RNA samples using a variety of gene expression microarray platforms and alternative gene expression profiling technologies across three independent laboratories, making the MAQC dataset unique both in its size and its content. The results showed that the median coefficient of variation for the within-laboratory replicates ranged from 5% to 15% for the various microarray platforms, whereas it was 10% to 20% for between-laboratory replicates. The results also reported an average 89% overlap of the differentially expressed genes between test sites using the same platform and a 74% overlap across different unicolour microarray platforms. Thereby, a high correlation was observed between the ranks of log ratios among the microarrays, indicating that all platforms, which use different approaches to measure gene expression, were detecting similar changes in gene abundance. In addition to the initial publication, Canales and colleagues compared the performance of the three quantitative gene expression technologies, such as quantitative RT-PCR approaches, to the five commercial microarray platforms based on the original MAQC dataset [35]. As expected they found a high level of correlation between the two types of platform, with only few discordant measurements. These were mainly explained by differences in probe sequence and by the fact that the microarray platforms showed a variable sensitivity to detect weakly expressed genes. Therefore, in this report, they highlighted the advantages of alternative gene expression profiling: assay specificity, lower detection threshold and an expanded assay range.

Alternatively, Patterson and colleagues compared one-colour versus two-colour microarrays to quantify mRNA abundance. Their results indicated that data quality is essentially equivalent between the two approaches and strongly suggest that this variable need not be a primary factor in decisions regarding experimental microarray design [36]. Also based on the original MAQC dataset and taking into consideration the measurement variability in the selection of differentially expressed genes using microarray data, some authors from the MAQC consortium developed a novel metrics called the Reproducibility Probability Score (RPS), a high RPS value meaning that the differential expression is more reproducible by another laboratory [37]. Herewith they wanted to convince researchers to introduce the inter-laboratory variability in identifying differentially expressed genes.

Although the first results of this MAQC huge collaborative effort support the use of gene expression profiling for basic and applied research, the work related to the first phase of the MAQC project did not examine other important parameters affecting the reproducibility of gene expression profiling studies such as sample collection and preservation, analytical and methodological approaches, as well data storage criteria. Clearly, sample preparation and storage are crucial issues for microarray experiment. Very soon, clinical investigators identified the need to assess the feasibility of using the RNA from fine needle biopsy samples in order for gene expression profiling to be applicable for clinical research or management of breast carcinoma. Several studies consistently reported that fine needle aspirations yielded enough material for DNA microarray analysis and observed similar gene expression profiles derived from FNAs and from core needle biopsy [38–40]. However, it should be noted that FNAs offer a purer representation of the tumour cell population, whereas transcriptional profiles from core needle biopsies include more representation from stroma elements. Therefore, the selection of the preferred needle biopsy sampling technique for genomic studies of breast carcinomas should depend on whether variable stroma gene expression is desirable in the samples.

Besides the type of sample considered for the analysis, sample preservation remains another critical parameter for microarray analyses. Although recent progress has been made with tumour RNA analysis in formalin-fixed paraffin-embedded (FFPE) tissue, the RNA fragmentation and chemical modifications that occurs with formalin fixation remain significant barriers to high-throughput DNA microarrays studies. In this view, the collection of high quality frozen specimens remains a crucial step. Indeed, RNA is inherently unstable and rapid changes may occur as a result of insults caused by tissue handling or ischemia. Several international groups and initiatives already addressed this issue. For example, Tubafrost[2], an initiative supported by the European Commission which aims to create an innovating virtual (networked) European human frozen tumour tissue bank for the whole scientific community, proposed standard operating procedures for the collection and storage of frozen specimens. The Fresh Tissue Working Group of the Breast International (BIG) and National Cancer Institute (NCI) also developed recently recommendations for fresh/frozen collection which were meant to set an identifiable standard for the collection of fresh/frozen tissue from BIG-auspiced and NCI-sponsored Group breast cancer trials[3]. These guidelines do not aim at defining unique procedures but rather allow flexibility in how frozen tissue is collected as long as the general

[2] www.tubafrost.org

[3] http://ctep.cancer.gov/forms/guidelines_fresh_tissue.pdf

principles of rapid collection and freezing are respected. An alternative to using fresh frozen samples is to perform gene expression profiling on formalin-fixed paraffin-embedded tissue, which is now possible using multiples real-time reverse transcriptase polymerase chain reaction (Q-RT-PCR). This is the strategy chosen by Paik and colleagues [19] to evaluate the Recurrence Score. Although this approach allows the measurement of mRNA levels of several genes, the use of FFPE tissue is still not suited for profiling thousands of genes. Therefore, one should first try to identify a consensus set of genes using high-throughput whole genome profiling that could then be validated using Q-RT-PCR and implemented to the clinic.

After data collection, the data are usually normalised to facilitate the comparison between different hybridisation assays, as normalisation should compensate for differences in labelling, hybridisation and detection methods. It should be emphasised that normalisation takes an often under-estimated critical role in the earlier stage of microarray data analysis since different normalisation procedures can lead to different expression data. There exist currently several approaches to normalise data [41–45]. The use of each of them depends on the type of array, the number of samples evaluated and the assumptions about biases in the data. Although several groups already tried to provide a comparison between different methods [46–48], the results are not consistent. In order to standardise the recording of data transformation and normalisation, the Microarray Gene Expression Data (MGED) Society[4], which was founded in 1999 as a collaborative effort regrouping many of the major microarray users and developers, launched the 'Data Transformation and Normalisation Working Group'[5] aiming at facilitating the sharing of microarray data. Their ultimate goal is not to oblige the microarray users to transform and normalise their data in a specific way, but instead to record accurately the way they performed it.

After the appropriate data have been normalised and possibly filtered, a variety of approaches may be chosen for further analysis. However, as highlighted by many bio-informaticians, many analyses suffer biases which mainly arise from the problem of multiple testing and/or inappropriate validations of the results. As an example, Dupuy and colleagues recently conducted a critical review of published microarray studies for cancer outcome [49]. Based on their observations, they developed new guidelines which could serve as an accessible and common basis for discussion among all cancer researchers involved in microarray investigations. In this context, the MACQ consortium now aims in the second phase of its project to understand the behaviour of various prediction rules and gene selection methods that may be applied to microarray data sets to generate clinical outcome predictors. It is anticipated that the results derived from this initiative, which should be available in autumn 2008, will help develop 'best practices' for the generation, analysis and application of microarray data in real-life settings.

Besides analytical biases, several methodological deficiencies may appear when reviewing published gene expression studies. The choice of end points for the identification of predictive gene signatures is one example. Many studies are using clinical measurements as a parameter for treatment response. However, the concordance between clinical and pathological response remains moderate. Thereby, it has been shown that pathological complete response might still not be the perfect surrogate for overall survival [50]. Furthermore, the choice of the regimen to be studied (combination chemotherapy as opposed to single agent) has a strong impact on the predictor which will be developed. We have seen that there clearly exists different molecular breast cancer subgroups, therefore evaluating a gene signature in the whole breast cancer population as opposed to a relevant molecular subgroup might be of crucial importance, as evaluating a predictor in an inappropriate cohort might lead to underestimation of its performance. Additionally, results are often based on sample sizes which are too small to draw meaningful results and unfortunately the sample size of biomarker studies is often driven by the availability of the tissues.

In order to improve and accelerate the process by which tumour markers that provide useful information for the management of cancer patients are adopted into clinical practice, the Statistics Committee of the NCI-EORTC Working Group on Cancer Diagnostics reported guidelines for tumour marker prognostic studies, which are the results of a collaborative effort among statisticians, clinicians and laboratory scientists. Although these guidelines are not specifically addressing the development of prognostic models from large numbers of candidate markers, such as those deriving from microarray studies, the recommendations regarding study design, pre-planned hypotheses, patient and specimen characteristics, assay and statistical methods remain also relevant to such studies. The Program for the Assessment of Clinical Cancer Tests[6] (PACCT) group of the US National Cancer Institute has now endorsed this effort.

Although many significant results have been derived from microarray studies, one limitation has been the

[4] www.mged.org

[5] http://smd-www.stanford.edu/mged/normalization.html

[6] http://www.cancerdiagnosis.nci.nih.gov/assessment/index.html

lack of standards for presenting and exchanging such data. In 2001, MGED developed the 'Minimum Information About a Microarray Experiment' (MIAME)[7], that describes the minimum information required to ensure that microarray data can be easily interpreted and that results derived from its analysis can be independently verified. The purpose of this checklist was to guide authors, journal editors and referees in helping them to ensure that the data supporting published results based on microarray experiments are made publicly available in a format that enables unambiguous interpretation of the data and potential verification of the conclusions. Several major scientific journals adopted MIAME recommendations as a requirement for publication of microarray experiments. In addition, MGED also strongly recommends that the data is made publicly available through one of the public repositories for microarray data. While some journals have already made this condition of acceptance, MGED feels that submission requirements should be applied consistently and that journals should recognise the Gene Expression Omnibus (GEO)[8] hosted by NCBI, Array Express[9] hosted by the European Bioinformatics Institute (EBI) and the Centre for Information Biology gene Expression Database (CIBEX)[10] as acceptable public repositories. As stated by Ball and colleagues, requiring authors to submit microarray data to a public database provides a number of considerable advantages to the entire research community: (1) it ensures that the published datasets will continue to be available into the future, (2) having the data available in a standardised format makes it more accessible and allows it to be integrated with other relevant data, (3) submission prior to publication makes it easier for referees to access the data confidentially, facilitating the review and publication process, and, (4) the standardisation of microarray data formats will enable the development of additional data analysis and integration tools and enabling scientists to access, query, and share data more easily [51]. The great potential of having different gene expression datasets available in the public domain was illustrated by the recent meta-analysis we carried out using publicly available gene expression data of ~3000 breast cancer patients [16], which would have been impossible if all these data would have been stored on local computers.

Clinical validation

A tumour marker is only clinically useful if the knowledge of its status promotes a change in practice that favourably affects clinical outcome. In the modified version of the 'Levels of Evidence Scale', Hayes and colleagues suggested categories to define the quality of the data which exists for a particular biomarker, with level V evidence being the weakest and level I the strongest one and considered as definitive [1].

Unfortunately, the majority of the gene expression studies never go further than level V or IV evidence, lacking further validation studies. However, for breast cancer, TRANSBIG, the translational research network founded by the Breast International Group, conducted an independent validation study of the 70 and 76-gene prognostic signatures in a series of 302 patients from five different centres, and across different statistical facilities. Although there was only a three gene-overlap between these signatures, both were validated in this patient cohort, even after adjustment for clinical risk [11,13]. Similarly, the RS was also validated in an independent population, moving these signatures to level III/IV evidence. According to Hayes' definitions, level I evidence can be obtained from a single, high-powered, prospective, randomised, controlled trial or from a meta-analysis or overview of multiple well-designed studies. Despite all the potentially useful gene signatures, the process of clinical validation to achieve level I evidence is just beginning for two of them, namely the 21 gene recurrence-score (ONCOTYPE DX- TAILORX) and the 70 gene Amsterdam signature (MAMMAPRINT®-MINDACT, see Fig. 1).

TAILORx® Trial Assigning IndividuaLised Options for Treatment (Rx), which is sponsored by the US NCI and coordinated by the Eastern Cooperative Oncology Group (ECOG), is the first study that has been developed as a result of the National Cancer Institute's PACCT. The study will enroll over 10,000 women at 900 sites in the USA and Canada. Women recently diagnosed with oestrogen-receptor and/or progesterone-receptor positive, Her2/neu-negative breast cancer that has not yet spread to the lymph nodes are eligible for this study. Patients with a low recurrence

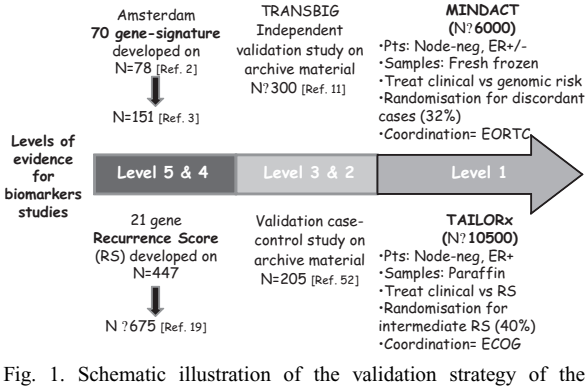

Fig. 1. Schematic illustration of the validation strategy of the 70-gene signature [2,3,11,52] and of the Recurrence Score [19].

[7] www.mged.org/Workgroups/MIAME/miame.html
[8] www.ncbi.nlm.nih.gov/geo/
[9] www.ebi.ac.uk/arrayexpress/
[10] http://cibex.nig.ac.jp/

score (RS) will receive standard hormonal therapy, patients with a high RS will receive combination chemotherapy followed by hormonal therapy, whereas patients presenting an intermediate RS (expected to represent 40% of the whole population) will be randomised to receive either a chemotherapy followed by hormonal therapy alone or hormonal therapy alone. Currently more than 1200 patients have already been included in TAILORx.

The Microarray for Node-negative Disease Avoids Chemotherapy (MINDACT) trial, which is partly sponsored by the European Commission, run under the auspices of the TRANSBIG consortium and co-ordinated by the European Organisation for Research and Treatment of Cancer (EORTC), will prospectively evaluate the added value of the 70-gene prognostic gene signature over the clinical criteria Adjuvant! Online software [www.adjuvantonline.com]. The primary objective of the MINDACT trial is to confirm that patients with a 'low risk' molecular prognosis and 'high risk' clinical prognosis can be safely spared chemotherapy without affecting clinical outcome. This trial just began recruiting patients. Besides their primary aim of evaluating the clinical applicability of these two signatures, these trials will also serve as an important resource for creating tissue and specimen banks to allow future evaluation of new molecular signatures and other technologies, such as proteomics, epigenomics, and pharmacogenomics, as they evolve.

Conclusions and perspectives

Currently there are only a few biomarkers with proven value in predicting prognosis or response to treatment for breast cancer patients (Her2/Neu and oestrogen receptors). Several gene expression studies have shown that a pattern of molecular markers might be more useful than individual markers and that these could have a tremendous potential. The fact that two gene expression signatures already entered the process of clinical validation is a clear sign that the use of expression profiles as biomarkers to predict disease prognosis and outcome is coming of age. Also, gene expression technology seems now to be mature enough to yield predictors that are reproducible within and across laboratories when the operating procedures are standardised. For example, huge initiatives such as MGED and MACQ are providing the microarray community with standardisation of data reporting, common analysis tools and useful controls that can help provide confidence in the consistency and reliability of these gene expression platforms.

However, a lot remains to be done, especially in identifying biomarkers predicting response to particular therapy as unpredictable efficacy and toxicity unfortunately still remain hallmarks of most anticancer therapies. Also, to date, most prognostic and predictive gene expression studies have focused on tumour characteristics, but it is likely that pharmacogenetics, genetic variability in the metabolism of therapeutic agents, as well as interactions between host and tumour cells also play an important role. Thereby, the application of the rapidly evolving high-throughput techniques for the realisation of genomic, proteomics and metabolomic profiles hold great promise to increase our biological understanding of the disease. But at the end, the final purpose of these different '-omic' approaches should not be to neglect the commonly used clinico-pathological markers but to try to find a way to make predictions more accurate by integrating different types of information and so provide the oncologists with a more accurate tool to facilitate the individual patient's treatment decision making.

Conflict of interest statement

None declared.

References

1 Hayes DF, Bast RC, Desch CE, *et al.* Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* 1996, **88**(20), 1456–1466.

2 van 't Veer LJ, Dai H, van de Vijver MJ, *et al*. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002, **415**(6871), 530–536.

3 van de Vijver MJ, He YD, van't Veer LJ, *et al*. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002, **347**(25), 1999–2009.

4 Wang Y, Klijn JG, Zhang Y, *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005, **365**(9460), 671–679.

5 Foekens JA, Atkins D, Zhang Y, *et al.* Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. *J Clin Oncol* 2006, **24**(11), 1665–1671.

6 Chang HY, Nuyten DS, Sneddon JB, *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci USA* 2005, **102**(10), 3738–3743.

7 Chang HY, Sneddon JB, Alizadeh AA, *et al.* Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* 2004, **2**(2), E7.

8 Glinsky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. *J Clin Invest* 2005, **115**(6), 1503–1521.

9 Liu R, Wang X, Chen GY, *et al.*The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* 2007, **356**(3), 217–226.

10 Sotiriou C, Wirapati P, Loi S, *et al.* Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006, **98**(4), 262–272.

11 Buyse M, Loi S, van't Veer L, *et al.* Multi-centre independent validation of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006, **98**(17), 1183–1192.

12 Ma XJ, Wang Z, Ryan PD, *et al.* A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* 2004, **5**(6), 607–616.

13 Desmedt C, Piette F, Loi S, *et al.* Strong time-dependency of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multi-centre independent validation. *Clin Can Res* 2007, **13**(11), 3207–3214.
14 Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 2006;355(6):560–569.
15 Bild AH, Potti A, Nevins JR. Linking oncogenic pathways with therapeutic opportunities. *Nat Rev Cancer* 2006, **6**(9), 735–741.
16 Sotiriou C, Haibe-Kains B, Desmedt C et al. Comprehensive molecular analysis of several prognostic signatures using molecular indices related to hallmarks of breast cancer: proliferation index appears to be the most significant component of all signatures. *Br Can Res Treat* 2006, **100**(suppl 1), 2020.
17 Loi S, Haibe-Kains B, Desmedt C, *et al.* Definition of clinically distinct molecular subtypes in estrogen receptor positive breast carcinoma through use of genomic grade. *J Clin Oncol* 2007, **25**(10), 1239–1246.
18 Jansen MP, Foekens JA, van Staveren IL, *et al*. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J Clin Oncol* 2005, **23**(4), 732–740.
19 Paik S, Shak S, Tang G, *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004, **351**(27), 2817–2826.
20 Jansen MP, Sieuwerts AM, Look MP, *et al.* HOXB13-to-IL17BR expression ratio is related with tumor aggressiveness and response to tamoxifen of recurrent breast cancer: a retrospective study. *J Clin Oncol* 2007, **25**(6), 662–668.
21 Ma XJ, Hilsenbeck SG, Wang W, *et al.* The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer. *J Clin Oncol* 2006, **24**(28), 4611–4619.
22 Folgueira MA, Carraro DM, Brentani H, *et al*. Gene expression profile associated with response to doxorubicin-based therapy in breast cancer. *Clin Cancer Res* 2005, **11**(20), 7434–7443.
23 Hannemann J, Oosterkamp HM, Bosch CA, *et al*. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2005, **23**(15), 3331–3342.
24 Chang JC, Wooten EC, Tsimelzon A, *et al.* Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients. *J Clin Oncol* 2005, **23**(6), 1169–1177.
25 Iwao-Koizumi K, Matoba R, Ueno N, *et al.* Prediction of docetaxel response in human breast cancer by gene expression profiling. *J Clin Oncol* 2005, **23**(3), 422–431.
26 Bertucci F, Finetti P, Rougemont J, *et al*. Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. *Cancer Res* 2004, **64**(23), 8558–8565.
27 Ayers M, Symmans WF, Stec J, *et al.* Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 2004, **22**(12), 2284–2293.
28 Potti A, Dressman HK, Bild A, *et al.* Genomic signatures to guide the use of chemotherapeutics. *Nat Med* 2006, **12**(11), 1294–1300.
29 Tan PK, Downey TJ, Spitznagel Jr EL, *et al*. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 2003, **31**(19), 5676–5684.
30 Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. *Nat Methods* 2005, **2**(5), 337–344.
31 Lee JK, Bussey KJ, Gwadry FG, *et al.* Comparing cDNA and oligonucleotide array data: concordance of gene expression across platforms for the NCI-60 cancer cells. *Genome Biol* 2003, **4**(12), R82.
32 Mecham BH, Klus GT, Strovel J, *et al.* Sequence-matched probes produce increased cross-platform consistency and more reproducible biological results in microarray based gene expression measurements. *Nucleic Acids Res* 2004, **32**(9), e74.
33 Barnes M, Freudenberg J, Thompson S, Aronow B, Pavlidis P. Experimental comparison and cross-validation of the Affymetrix and Illumina gene expression analysis platforms. *Nucleic Acids Res* 2005, **33**(18), 5914–5923.
34 MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006, **24**(9), 1151–1161.
35 Canales RD, Luo Y, Willey JC, *et al.* Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotechnol* 2006, **24**(9), 1115–1122.
36 Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, *et al.* Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. *Nat Biotechnol* 2006, **24**(9), 1140–1150.
37 Lin G, He X, Ji H, Shi L, Davis RW, Zhong G. Reproducibility Probability Score- incorporating measurement variability across laboratories for gene selection. *Nat Biotechnol* 2006, **24**(12), 1476–1477.
38 Sotiriou C, Khanna C, Jazaeri AA, Petersen D, Liu ET. Core biopsies can be used to distinguish differences in expression profiling by cDNA microarrays. *J Mol Diagn* 2002, **4**(1), 30–36.
39 Assersohn L, Gangi L, Zhao Y, *et al.* The feasibility of using fine needle aspiration from primary breast cancers for cDNA microarray analyses. *Clin Cancer Res* 2002, **8**(3), 794–801.
40 Symmans WF, Ayers M, Clark EA, *et al.* Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma. *Cancer* 2003, **97**(12), 2960–2971.
41 Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003, **31**(4), e15.
42 Schadt EE, Li C, Ellis B, Wong WH. Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. *J Cell Biochem Suppl* 2001, **37**, 120–125.
43 Quackenbush J. Microarray data normalization and transformation. *Nat Genet* 2002, **32**(Suppl), 496–501.
44 Yang IV, Chen E, Hasseman JP, *et al.* Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biol* 2002, **3**(11), research0062.
45 Yang YH, Dudoit S, Luu P, *et al.* Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 2002, **30**(4), e15.
46 Shedden K, Chen W, Kuick R, *et al.* Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC Bioinformatics* 2005, **6**, 26.
47 Ploner A, Miller LD, Hall P, *et al.* Correlation test to assess low-level processing of high-density oligonucleotide microarray data. *BMC Bioinformatics* 2005, **6**, 80.
48 Park PJ, Cao YA, Lee SY, *et al.* Current issues for DNA microarrays: platform comparison, double linear amplification, and universal RNA reference. *J Biotechnol* 2004, **112**(3), 225–245.
49 Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst* 2007, **99**(2), 147–157.
50 Fisher B, Bryant J, Wolmark N, *et al.* Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. *J Clin Oncol* 1998, **16**(8), 2672–2685.
51 Ball CA, Brazma A, Causton H, *et al.* Submission of microarray data to public repositories. *PLoS Biol* 2004, **2**(9), E317.
52 Habel LA, Shak S, Jacobs MK, *et al.* A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Res* 2006, **8**(3): R25.