International Journal of
Surgical Pathology
Volume 17 Number 4
August 2009 285-302

10.1177/1066896908328577
http://ijsp.sagepub.com

Microarray-based Gene Expression Profiling as a Clinical Tool for Breast Cancer Management: Are We There Yet?

Felipe Correa Geyer, MD, and Jorge Sergio Reis-Filho, MD, PhD, FRCPath

Breast cancer is a heterogeneous disease, encompassing several histological types and clinical behaviors. Current histopathological classification systems are based on descriptive entities with prognostic significance. Few prognostic and predictive markers beyond those offered by histopathological analysis are available. High-throughput molecular technologies are reshaping our understanding of breast cancer, of which microarray-based gene expression has received most attention. This method has been used to derive a molecular taxonomy for breast cancer, which has provided interesting insights into the biology of the disease. Class prediction studies have generated a multitude of prognostic/predictive signatures, which herald the promise for an improvement in treatment decision making. However, most of the signatures developed to date seem to have discriminatory power almost restricted to estrogen receptor–positive disease. This review addresses the contribution of gene expression profiling to our understanding of breast cancer and its clinical management.

Keywords: gene expression; microarray; gene signature; prognostic; predictive

Breast cancer is a heterogeneous disease encompassing a multitude of pathological entities that are reported to have distinct clinical behaviors.[1,2] The complexity of breast cancer has long been acknowledged by histopathologists, who have endeavored to devise classification systems to account for its diversity. However, current classification systems are largely descriptive and are based on morphological entities that have been shown to harbor prognostic implications usually on retrospective analysis. With the boom of high-throughput technologies and the availability of humanized antibodies against known oncogenes and hundreds of small molecule inhibitors (most of which are still confined to the realm of preclinical studies), novel challenges have emerged. In particular, pathologists have been asked to provide additional predictive information and to develop robust and reliable companion diagnostics. It should not come as a surprise, therefore, that for the success of targeted therapies and individualized medicine, a predictive rather than purely prognostic classification system is required.

Despite the translational research efforts of the 1980s and 1990s, which were largely focused on the study of single proteins or genes at a time, only 3 biomarkers have become part of the armamentarium of breast pathologists and are routinely used to determine breast cancer therapy: estrogen receptor (ER), progesterone receptor (PR), and HER2. Interestingly, all of these biomarkers have optimal negative predictive values (ie, patients with ER-negative breast cancer are highly unlikely to respond to endocrine therapy; HER2-negative breast cancers fail to respond to humanized monoclonal antibodies against HER2). However, the positive predictive value of these biomarkers is rather limited, with a

From the Molecular Pathology Laboratory, Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London, UK.

Address correspondence to: Jorge S. Reis-Filho, MD, PhD, FRCPath, The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, 237 Fulham Road, London, UK SW3 6JB; e-mail: Jorge.Reis-Filho@icr.ac.uk.

substantial proportion of patients with HER2-positive disease either harboring de novo resistance or developing resistance to trastuzumab over time.[3] Therefore, new biomarkers and therapeutic targets are urgently needed to improve breast cancer patient management.

The entire field of cancer taxonomy, biomarker development, and identification of prognostic and predictive markers started to change with the advent of high-throughput technologies, of which microarray-based gene expression profiling has been the most extensively used. A molecular taxonomy pioneered by the Stanford group[4,5] has been reasonably reproducible in different cohorts.[6,7] Although this taxonomy should be perceived as a draft, it has been progressively adopted as a working model for a molecular classification by independent groups. Several prognostic and/or predictive signatures (collection of genes that taken together can classify tumors in different prognostic/predictive groups) have been developed.[8-24] Not surprisingly, these signatures have attracted enormous attention, given that they are usually described as capable of predicting relapse more accurately than algorithms based on traditional clinicopathological parameters (eg, Adjuvant!Online). Some scientists and clinicians have even described current pathology methods as unsophisticated and obsolete and compared them to ritual practices by ancient tribes.[25]

The contribution of gene expression profiling to our understanding of breast cancer biology is undeniable, and it is becoming increasingly evident that this approach may lead to the identification of biomarkers and mechanisms of response/resistance to therapy.[26-28] One could argue, however, that the main contribution of microarrays so far has been the rediscovery of breast cancer heterogeneity and the opportunity to devise a systematic way of classifying breast cancer into biologically and clinically meaningful subgroups. In fact, the impact of these findings has been such that clinical trials are being designed in ways that can account for this diversity,[4,5] and therapeutic approaches are being tested in specific subgroups rather than the entire population of breast cancer patients.[29] Therefore, it should not come as a surprise if in the not-so-distant future, microarray-based expression profiling, or most likely one of its derivatives, is incorporated into breast pathology practice. However, several issues must be resolved before this technique can be translated to routine clinical use. The initial enthusiasm has been tempered recently by critical analyses of the conclusions we can draw from these studies.[30-33] Furthermore, meta-analyses of microarray signatures have demonstrated that we have not progressed much beyond ER, PR, HER2, and proliferation.[34] In this review we address the contribution of gene expression profiling to breast cancer management.

Molecular Portraits of Breast Cancer: Getting Beyond Pretty Pictures

Microarray-based gene expression profiling studies can be performed in multiple ways. The initial approaches focused on the use of unsupervised methods to test whether microarrays would provide biologically and/or clinically meaningful information about breast cancer diversity. These studies, pioneered by the Stanford group,[4,5] were largely based on hierarchical clustering analysis of relatively small cohorts of breast cancers. Unsupervised hierarchical clustering using an "intrinsic gene list" led to the identification of 5 molecular groups according to their expression pattern, namely luminal A, luminal B, HER2, basal-like, and normal breast-like (Figure 1).

The most obvious distinction observed by microarray analysis is between the transcriptomes of ER-positive and ER-negative breast cancers (Figure 1). Luminal tumors are described as those that show expression patterns reminiscent of normal luminal epithelial cells of the mammary gland, including consistent expression of low molecular weight cytokeratins 8/18, ER, and genes associated with an active ER pathway.[4,5,7,35,36] At least 2 subgroups of luminal tumors have been identified: luminal A tumors, which are usually of low histological grade, have an excellent prognosis, and show high levels of expression of ER-activated genes; and luminal B tumors, which are more often of higher histological grade and have higher proliferation rates and a significantly worse prognosis than luminal A tumors.[4,5,7,35,36] Normal breast-like cancers are rather poorly characterized; one of the defining features of these tumors is that they consistently cluster together with samples of fibroadenomas and normal breast. The clinical significance of normal breast-like tumors is yet to be determined,[4,5,7,35,36] and some have suggested that this subgroup may be a mere artifact of expression profiling (ie, disproportionally high content of stromal cells), given that identification of this group of tumors by microarrays is less stable

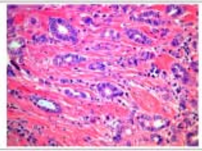
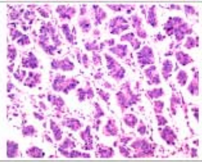
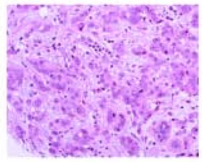
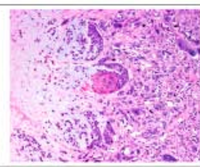
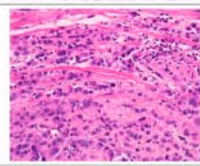
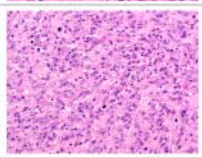
Molecular Subgroup	ER	PR	HER2	Basal Markers	Other Markers	Proliferation	Molecular Pathways	Histological type°	Example
Luminal A	+++	+++	-	-		Low	Luminal CKs; ER-responsive genes (e.g. *CCND1*, *TFF1*, *GATA3*)	Tubular carcinoma, Cribriform carcinoma, Low grade IDC-NST, Classic Lobular carcinoma	
Luminal B	+	+/-	+/-	+/-		High	Luminal CKs; ER-responsive genes (e.g. *CCND1*, *TFF1*, *GATA3*); *TP53* mutations	IDC-NST, Micropapillary carcinoma	
HER2/ER-	-	-	+	+/-		High	HER2 amplicon genes (e.g. *GRB7*, *TOPO2A*, *GATA4*); *TP53* mutations	High grade IDC-NST	
Basal-like	-	-	-	+		High	Basal CKs; *TP53* mutations; DNA repair defects (BRCA1 pathway)	High-grade IDC-NST, Metaplastic carcinoma, Medullary carcinoma, Adenoid Cystic carcinoma	
Molecular apocrine	-	-	+/-	+/-	AR+	High	Androgen signalling; Fatty acid and lipid synthesis	Apocrine carcinoma, IDC-NST, Pleomorphic Lobular carcinoma	
Claudin-low	-	-	-	+/-	E-cad-*	High	Cell-cell adhesion (*Claudins 3, 4, 7, Occludin, E-cadherin*); mesenchymal and lymphocyte-genes	Grade 3 IDC-NST, Metaplastic carcinoma	

Figure 1. Molecular subgroups of breast cancer. °Most suitable. *Not lobular morphology. Basal markers: EGFR, CKs 5/6, 14, and 17. ER, estrogen receptor; PR, progesterone receptor; AR, androgen receptor; CK, cytokeratin; IDC-NST, E-cad, e-cadherin invasive ductal carcinoma of no special type.

when fine needle aspiration biopsies are used[37] or when microdissected samples are subjected to expression array analysis (J. S. Reis-Filho and R. Natrajan, unpublished observations). HER2 tumors are usually ER negative and characterized by overexpression of HER2 and genes associated with HER2 pathway and/or *HER2* amplicon on 17q12. HER2 cancers have very aggressive clinical behavior; however, they are amenable to novel tailored therapies using either humanized monoclonal antibodies against HER2 or HER2 tyrosine kinase inhibitors.[4,5,7,35,36,38-42] Although the vast majority (>80%) of HER2 cancers as defined by microarrays harbor *HER2* gene amplification or HER2 3+ immunohistochemical expression,[35,37] not all tumors that are *HER2* amplified fall into the HER2 cluster by expression arrays analysis. There is also evidence to suggest that some *HER2*-amplified, ER-positive cancers fall within the luminal B subtype rather than the HER2-microarray subtype.[35] Basal-like cancers, another group of predominantly ER-negative cancers, are so named because the neoplastic cells of this tumor type consistently express genes usually found in normal basal/myoepithelial cells of the breast, including high molecular weight cytokeratins (5/6, 14, and 17), P-cadherin, caveolins 1 and 2,[43-50] nestin,[51] and epidermal growth factor receptor (EGFR)[45] and, in a minority of cases, harbor *EGFR* gene amplification[36] or aneusomy.[52] These tumors are usually of high histological grade and characterized by high mitotic indices, the presence of central necrotic zones, pushing borders, conspicuous lymphocytic infiltrate, and typical/atypical medullary features.[53,54] In addition, metaplastic elements are not uncommon.[55] The morphological and immunohistochemical features of basal-like cancers are remarkably similar to those described for tumors arising in *BRCA1* germline mutation carriers,[49] and there is a growing body of evidence to suggest that BRCA1 pathway is dysfunctional in sporadic basal-like cancers.[56] In fact, engineered mouse models targeting *Brca1* and *Trp*53 genes in luminal or basal cells of the mouse mammary gland resulted in the development of tumors displaying morphological and immunohistochemical features that recapitulate those seen in human basal-like breast carcinomas.[57,58]

Although basal-like breast cancers have only made it to the forefront of breast cancer research in the last decade after their rediscovery by microarray analysis, the existence of a subgroup of aggressive breast carcinomas harboring features of myoepithelial/basal differentiation was first documented more than 40 years ago,[44,59-62] and studies demonstrating that these tumors more often express the immunohistochemical features now known to be characteristic of basal-like cancers were first published in the last decade.[63] Basal-like breast cancers are typically ER negative, PR negative, and HER2 negative (triple-negative); however, the terms *basal-like* and *triple-negative* are not synonymous. Although the vast majority of triple-negative cancers are of basal-like phenotype[64-66] and the vast majority of tumors expressing "basal" markers are triple-negative,[45,67-70] there is a significant number of triple-negative cancers that do not express basal markers and a small but still significant subgroup of basal-like cancers that express either hormone receptors or HER2.[37,65,67,69-72] Bertucci et al[67] addressed this issue directly and confirmed that not all triple-negative tumors when analyzed by gene expression profiling were classified as basal-like cancers (ie, only 71% were of basal-like phenotype) and not all basal-like breast carcinomas classified by expression arrays displayed a triple-negative phenotype (ie, 77% were of triple-negative phenotype). Taken together, these results are in accord with the concept that the triple-negative phenotype is not an ideal surrogate marker for basal-like breast cancers.[66,67,73] Furthermore, there are several lines of evidence to suggest that the group of triple-negative cancers is substantially more heterogeneous than the group encompassed by basal-like breast cancers.[66-68] This concept is particularly important when results of clinical trials on triple-negative cancers are interpreted[67]; we would argue that a subgroup analysis of the core basal subgroup of triple-negative cancers (ie, triple-negative cancers that express basal markers) should be systematically performed. Basal marker–positive triple-negative cancers have been shown in retrospective studies to have a more aggressive clinical behavior[68] and to have more consistent dysfunction of BRCA1 pathway than triple-negative, basal marker–negative cancers.[49,56]

Although this molecular classification has captured the attention of the scientific and medical communities, its introduction in the clinic is still a work in progress. Surrogate markers for the intrinsic gene list molecular taxonomy are being developed. Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)–based methods have already been generated and validated in archival, formalin-fixed, paraffin-embedded tissue samples.[74-76] Nielsen et al[45] described an immunohistochemical panel of 4 markers (ER, HER2, EGFR, and cytokeratins 5/6) that identifies basal-like cancers with 100% specificity and 76% sensitivity and has been further confirmed to be of prognostic significance.[68] However, other clinically meaningful subgroups may emerge from further class discovery studies. For instance, a "molecular apocrine" subgroup, which shares some features with the HER2 subtype,[77] and a "claudin-low" subgroup, which comprises tumors that have transcriptomic features suggestive of a stem cell-like phenotype, have been recently described[78] (Figure 1). However, whether claudin-low tumors constitute a separate molecular group of breast cancers or are a mere subgroup of basal-like cancers remains to be determined.

Class discovery studies performed to date have preferentially analyzed samples of invasive ductal carcinomas of no special type (IDC-NST) and a limited number of invasive lobular carcinomas.[4,5] Until very recently, special types of breast cancer had largely been systematically ignored. This issue was recently addressed by Weigelt et al,[79] who performed a detailed expression profiling analysis of special types of breast cancer and demonstrated that some special types of breast cancer pertain to a single molecular subgroup. For example, medullary and adenoid cystic carcinomas were consistently of basal-like phenotype. Micropapillary carcinomas clustered together in a distinct branch within the ER-positive tumors, confirming that this histological type is a distinct entity, a feature that is also supported by genetic analysis based on array comparative genomic hybridization.[80] Tubular carcinomas showed remarkable similarities

Table 1. Correlation Between Molecular Subtypes and Histological Special Types[77–79,83]

Histological Special Type	Luminal	HER2+/ER–	Basal-like	Molecular Apocrine	Claudin-low
Tubular carcinoma	X				
Cribriform carcinoma	X				
Mucinous carcinoma	X				
Micropapillary carcinoma	X				
Neuroendocrine carcinoma	X				
IDC osteoclast	X				
Classic ILC	X			X*	
Pleomorphic ILC	X	X		X	
Apocrine carcinoma	X	X		X	
Adenoid cystic carcinoma			X		
Medullary carcinoma			X		
Metaplastic carcinoma			X		X†

NOTES: ER = estrogen receptor; IDC osteoclast = invasive ductal carcinoma with osteoclastic giant cells; ILC = invasive lobular carcinoma.
*A minority. †Hypothetical.

to lobular carcinomas at the transcriptome level, supporting the concept of a "low-grade breast neoplasia family" previously described.[81,82] A novel neuroendocrine cluster composed by neuroendocrine and mucinous carcinomas has been found; its significance, however, remains to be determined. Metaplastic carcinomas have consistently been shown to fall into the basal-like cluster by immunohistochemical[55] and gene expression profiling studies[79,83] (Table 1). However, a recent study undertaking pathway analysis revealed that PTEN, topoisomerase IIα, and DNA repair pathways, including BRCA1 pathway, are significantly more down-regulated in metaplastic carcinomas compared with basal-like IDCs, which may at least in part explain their reported poor responses to chemotherapy.[83]

Although not yet comprehensively studied by gene expression profiling, a special type of cancer that is an obvious example of genotypic-phenotypic correlation in breast cancer is the secretory carcinoma. These tumors have been shown to consistently harbor a t(13;15)(p13;q24) chromosomal translocation,[84] involving the genes *ETV6* and *NTRK3*.[84-87] Subsequent studies have shown that this translocation in breast cancer is specific to secretory breast carcinomas and that even tumors once thought to be variants of this special type (ie, acinic cell carcinomas of the breast) do not harbor this translocation.[88]

As expected, most special types have been shown to be more homogeneous at the molecular level than IDC-NSTs, and some do constitute different entities. Further large-scale studies that include all histological types and preferentially incorporate gene expression and genetic analyses are still needed to comprehensively evaluate the molecular features of breast cancer. However, it is anticipated that by combining morphological, immunohistochemical, and molecular techniques, a more biologically and clinically meaningful classification of breast cancer, similar to that adopted for the diagnosis and management of leukemias and lymphomas, will emerge.

Prognostic Signatures: Sound Improvement or Noisy Overstatement?

Adjuvant!Online (www.adjuvantonline.com), Nottingham Prognostic Index,[89] St Gallen guidelines,[90] and National Institutes of Health (NIH) criteria[91] are the tools and algorithms often used by oncologists to decide which patients should receive chemotherapy. These algorithms are based on clinicopathological parameters including tumor size, histological grade, ER status, vascular invasion, and presence of lymph node metastasis. However, oncologists, pathologists, and surgeons have already come to terms with the concept that these parameters are not sufficient to capture the complexity of breast cancer and to tailor therapy for individual patients. Furthermore, the actual proportion of patients who actually benefit from adjuvant chemotherapy regimens is rather limited.

Although numerous studies using a candidate approach have identified potential prognostic markers (eg, p53, Mib1, MUC1, to name but a few), none of these has been translated into clinical practice. Given the need for better prognostic factors for breast cancer, it is not surprising that microarrays have been tested as a method to identify novel prognostic factors. Studies designed to devise gene expression–based multivariate function (referred to as predictors or gene signatures) are called *class prediction analyses*. The aim of this type of studies is to devise a gene signature that accurately predicts the class membership of a new sample on the basis of the expression levels of key genes. Such predictors can be used for many types of clinical decision making, including diagnosis, prediction, and, most often, prognostication. Using this supervised approach, investigators have developed numerous microarray-based prognostic gene signatures.[8-17,20-24] Unlike conventional immunohistochemical prognostic markers, these are multigene predictors of outcome. As with any type of multidimensional data where multiple comparisons are made, the risk of data overfitting is high, and several of the prognostic signatures have been shown not to perform as well in validation series, when compared with the original data sets from which they were derived.

The first breast cancer prognostic signature described was the 70-gene signature from the Amsterdam group.[22] This classifier was derived from the analysis of 78 frozen samples from small (<5 cm) lymph node–negative breast cancers from patients younger than 55 years treated at the Netherlands Cancer Institute. By comparing the expression profiles of tumors from patients who developed distant metastasis within 5 years with those who did not, the authors identified a multigene prognostic signature comprising 70 genes. This signature could accurately identify the patients who did not develop metastasis within 5 years (good prognosis) and those who relapsed within 5 years (poor prognosis). In subsequent studies, the signature was validated in larger cohorts, with node-positive and node-negative patients, and statistical analyses have shown that it outperforms traditional systems based on clinical and histological parameters in predicting disease outcome.[92-95] The major strength of the signature is its high accuracy in identifying low-risk patients who could be spared adjuvant aggressive chemotherapy. However, this signature has a negligible discriminatory power in ER-negative breast cancers, with <4% of patients with ER-negative breast cancer being considered as having good prognosis. Despite the multiple issues raised in relation to the applicability of the 70-gene signature (aka, MammaPrint) and to its limited value in ER-negative disease, this predictor is currently being tested in the MINDACT (Microarray In Node-negative Disease may Avoid ChemoTherapy), a prospective randomized study comparing the 70-gene signature with the common clinicopathological criteria in selecting patients for adjuvant chemotherapy in node-negative breast cancer. The design of this clinical trial has been subjected to numerous amendments; in its latest format, it will require 6000 patients, and only patients with discordant results between the 70-gene signature and clinicopathological parameters will be randomized.[96] Although this trial has the potential to settle once and for all the validity of this prognostic signature, it will take years to be completed and it is possible that by the time of its completion, this approach will have already been rendered obsolete. In fact, 13 months after the activation of the trial, only 265 participants had been accrued (http://www.breastinternationalgroup.org/BIG/Clinical_Trials/Active_trials/BIG__304_MINDACT_EORTC_10041.aspx).

A slightly different approach has been undertaken by Wang et al,[23] who first identified 16 and 60 genes associated with relapse in ER-positive and ER-negative breast carcinomas, respectively, and assembled them into a 76-gene prognostic signature. Multicenter studies have further validated this predictor in larger cohorts.[97,98] Like the 70-gene signature, this classifier was more accurate than the St Gallen[90] and NIH[91] methods. However, it has been observed that this 76-gene predictor is strongly time dependent[97] and that its 16-gene component had no predictive power when applied to a series of 71 ER-negative patients.[99]

Chang et al[9] studied the genomic response of fibroblasts to serum exposure and created the wound-response signature. These authors observed that the expression of the genes belonging to the signature in a proportion of breast, gastric, and lung carcinomas predicted poor overall survival and increased risk of metastasis. The authors then validated their findings using the database of 295 breast cancer patients that had previously been used to create and validate the 70-gene signature.[100] The wound-response signature could identify 90% of the patients who developed metastasis and could theoretically have spared 30% of women, who did not develop metastasis, from chemotherapy.[100]

In breast cancer, cells with high expression of CD44 and no or undetectable expression of CD24 have been purported to have high tumorigenic capacity.[101] Liu et al,[14] driven by the growing body of evidence which suggests that only a minority population of cancer cells within a tumor are actually tumorigenic, compared the gene expression profile of $CD44^{+}$/$CD24^{-/low}$ breast cancer cells with that of normal breast epithelium and generated a 186-gene "invasiveness" signature. This signature was an independent prognostic factor and, again, outperformed St Gallen[90] and NIH[91] criteria in identifying patients who would not benefit from adjuvant systemic therapy. Interestingly, the invasiveness gene signature was also associated with prognosis in medulloblastoma, lung cancer, and prostate cancer.[14]

The promise for improvement in treatment decision making attracted great attention to these signatures. Several additional classifiers have been published and are currently being developed (Table 2). However, many doubts regarding the reliability and reproducibility of this approach have arisen.[33,36,102] Subsequent statistical analyses have called into question the actual contribution of these signatures to breast cancer patient management.[30,31,103] One of the reasons for this apparent failure of microarrays in realizing their potential stems from the disparate paces of microarray technology development and the development of bioinformatics and statistics applied to microarray analysis. For instance, in the beginning of this century, data overfitting was a poorly understood concept and methods for power calculation for microarray analysis were yet to be described.[104] Fortunately, this field is maturing very rapidly and there are clear guidelines as to how a therapeutically significant gene signature should be developed and validated.[105,106] Unfortunately, none of the microarray-based signatures described to date fulfills all criteria required. In addition, although microarrays are quite reproducible and can certainly be applied to class discovery studies and preclinical analysis, their accuracy and reproducibility may not be sufficient for clinical use. For instance, up to 15% of error in qualitative assessment of gene expression has been demonstrated when the optimal samples were analyzed with the same platform and protocols in different laboratories.[107] Furthermore, the applicability of microarrays to readily available, formalin-fixed, paraffin-embedded material is limited, given that this technology has been shown to perform suboptimally when RNA extracted from this type of sample is used.[108] Although Illumina (http://www.illumina.com/) has recently provided a method for microarray-based gene expression analysis of formalin-fixed paraffin-embedded samples (DASL gene expression), which has been shown to produce reasonable results when applied to matched frozen and formalin-fixed paraffin-embedded specimens,[109] the above signatures are yet to be converted into a platform that is compatible with those samples.

Given the difficulty of implementing RNA-preserving tissue collection in the clinical practice, qRT-PCR-based signatures[18] and immunohistochemical panels[19] suitable for the analysis of formalin-fixed paraffin-embedded samples have been developed as alternative methods for expression profiling analysis. The prototype of this approach is Oncotype DX,[18] which has been developed to be applied in early-stage, endocrine-responsive, tamoxifen-treated patients. The assay is based on the messenger RNA (mRNA) expression levels of 21 genes (16 cancer-related genes and 5 reference genes). These 16 genes comprise components of ER pathway, proliferation, *HER2* amplicon, and invasion. The expression levels of these genes as defined by qRT-PCR are subjected to a mathematical algorithm and provide the recurrence score. Patients with a high recurrence score in material from previously performed clinical trials have been shown to benefit from chemotherapy. Intermediate and low scores are assembled under the good prognosis group, and chemotherapy would be of limited value. The results have been validated in different cohorts,[110,111] including node-positive patients.[112] Although some have argued that this predictor only recapitulates the information provided by grade, ER, HER2, and MIB1,[113] that the influence of the proportion of stromal cells in the final result of the test needs to be thoroughly examined, and that the reproducibility of this assay in a large cohort of samples is yet to be fully characterized, this predictor has gained wide acceptance by the oncology community, and major medical insurance providers in North America have already adopted Oncotype DX as a means to identify which patients with ER-positive cancers should receive chemotherapy. According to the Genomic Health Web site, to date 7500 physicians have ordered the Oncotype DX for more than 65 000 patients. As the 70-gene signature, it has reached the final step in prospective testing. The TAILORx (Trial Assigning Individualized Options for Treatment [Rx]) trial, which is currently being run in the United States and is led by the North American Breast Cancer Intergroup, has been designed to test the clinical impact of Oncotype DX in a prospective fashion.[114,115] The trial developers plan to enroll more than 10 000[115] node-negative, ER-positive, HER2-negative patients. Patients with high recurrence scores will receive hormone therapy and chemotherapy,

Table 2. Summary of Gene Signatures and Their Application in Breast Cancer Patient Management

Signature	Biological Hypothesis	Method	Platform	Sample Conditions	No. of Genes	Indication*	Commercial Status/ Manufacturer	References
70-gene/ MammaPrint	Clinical outcome	Microarray	Agilent (oligonucleotides)	Fresh frozen	70	ER+	On the market/ Agendia BV	22, 92–96
76-gene	Clinical outcome	Microarray	Affimetrix (oligonucleotides)	Fresh frozen	76	ER+	Not commercially available	23, 97–99
Wound response	Wound healing	Microarray	cDNA (custom made)	Fresh frozen	512	ER+	Not commercially available	9, 88
Invasiveness gene signature	Tumorigenic cells CD44+/ CD24–/low	Microarray	Affimetrix (oligonucleotides)	Fresh frozen	186	ER+	Not commercially available	14
Recurrence score/ Oncotype DX	Clinical outcome	qRT-PCR	TaqMan	FFPE	21	ER+	On the market/ Genomic Health	18, 110–114, 116
Mammostrat	Clinical outcome	IHC	IHC	FFPE	5	ER+	On the market/Applied Genomics	19
Genetic grade index	Histological grade	Microarray	Affimetrix (oligonucleotides)	Fresh frozen	97	ER+	Not commercially available	20, 123
5-gene molecular grade index	Histological grade	Microarray	Affimetrix (oligonucleotides)/ TaqMan	Fresh frozen/ FFPE†	5	ER+	Not commercially available	16
6-gene grade signature	Histological grade	Microarray	Affimetrix (oligonucleotides)	Fresh frozen	6	ER+	Not commercially available	12
2-gene ratio	Response to tamoxifen	Microarray	Agilent (oligonucleotides)/ TaqMan†	Fresh frozen/ FFPE†	2	ER+	Not commercially available	16, 17, 127–129
181-gene signature	Response to tamoxifen	Microarray	Affimetrix (oligonucleotides)	Fresh frozen	181	ER+	Not commercially available	15
Pharmacogenomic predictor	Response to chemotherapy	Microarray	Affimetrix (oligonucleotides)	Fresh frozen	26	ER+ and ER–	Not commercially available	11
DTF/SFT signature	Fibroblastic reaction patterns	Microarray	cDNA	Fresh frozen	786	ER+	Not commercially available	24, 141
Stroma-derived prognostic predictor	Epithelial-mesenchymal interaction	Microarray	Agilent (oligonucleotides)	Fresh frozen	26	ER+ and ER–	Not commercially available	10
Extracellular matrix signature	Epithelial-mesenchymal interaction	Microarray	Agilent (oligonucleotides)	Fresh frozen	278	ER+	Not commercially available	8
7-gene immune response module	Clinical outcome	Microarray	Multiple	Fresh frozen	7	ER–	Not commercially available	21, 155

NOTES: ER = estrogen receptor; cDNA = complementary DNA; qRT-PCR = quantitative reverse transcriptase polymerase chain reaction; FFPE = formalin-fixed paraffin-embedded; IHC = immunohistochemistry; DTF = desmoids-type fibromatosis; SFT = solitary fibrous tumor.
* Most applicable. †Derived from microarray data and later validated in a qRT-PCR format.

whereas patients with low scores will be treated with hormone therapy alone. Those with intermediate scores will be randomized to hormone therapy or hormone therapy plus chemotherapy. The results of this trial are eagerly awaited because a recent study showed that about half of the patients referred to undergo the test in real life were considered to be at intermediate risk.[116]

Genomic Grading: The End of Nottingham Grade?

It has been demonstrated that histological grade, more than any other clinicopathological parameter or tumor intrinsic characteristic, is associated with the type, pattern, and complexity of molecular changes seen in breast cancer and its precursors.[1,2] This concept stems not only from comparative genomic hybridization studies but also from microarray-based expression profiling analysis. Molecular genetic analysis has demonstrated remarkably distinct patterns of genetic aberrations in grade I and grade III breast cancers. In particular, deletions of 16q, the most prevalent numerical aberration in grade I breast cancers, is seen in <20% of grade III breast carcinomas, suggesting that progression from low-grade to high-grade breast cancer is uncommon.[2,117-119] Furthermore, Ma et al[120] analyzed matched normal, pre-invasive, and invasive breast lesions of different grades by means of expression profiling analysis and observed that samples preferentially clustered according to histological grade rather than stage.

Histological grading of breast cancer has long been used as an important parameter for treatment decision making. Currently, it is performed by histopathological analysis using the modified Bloom and Richardson method,[121] also known as Nottingham grade. Despite the importance of this grading system, concordance between pathologists has been reported to be suboptimal.[122] In fact, intra-observer and interobserver agreement ranges from 50% to 86%. Although the clinical significance of grade I and grade III is well defined, approximately 50% of the tumors are classified as of grade II, a grade group whose interobserver reproducibility is clearly suboptimal and whose clinical significance remains equivocal.

Sotiriou et al[20] have undertaken a hypothesis-driven approach with the aim of refining the current grading system and defining the molecular basis of histological grade. Histologically defined grade I and grade III ER-positive breast carcinomas were subjected to microarray-based gene expression profiling, and using a class prediction approach, the investigators devised a 97-gene expression grade index (GGI). The GGI identified breast cancers classified as histological grade I and grade III tumors in the validation set with an accuracy of ~90% and was more strongly associated with relapse-free survival than histological grade. Importantly, GGI could stratify histological grade II tumors into GGI low-grade and GGI high-grade categories and could show that these groups were of prognostic significance: histological grade II GGI low tumors had outcomes similar to GGI low cancers, whereas histological grade II GGI high tumors had outcomes similar to GGI high cancers. GGI was subsequently validated in an independent data set by the same group.[123]

Two independent groups used analogous approaches and achieved similar results.[12,16] Ivshina et al[12] and Ma et al[16] identified a 6 grade–associated gene signature and a 5-gene molecular grade index, respectively. The latter has been converted into a qRT-PCR assay suitable for formalin-fixed, paraffin-embedded, archival patient material, which may expedite its incorporation into clinical decision making for breast cancer patients.[16] It has also been proposed that grading signatures could replace histological grade in currently used prognostic models such as Adjuvant!Online and the Nottingham Prognostic Index.[12,124,125] Ivshina et al[12] described that when the 6-gene grade signature rather than histological grade was used in the Nottingham Prognostic Index, some patients classified as having poor prognosis shifted to the good prognosis group. Hence, the use of microarray-based genomic signatures to assess tumor grade may improve the identification of patients who could be spared toxic adjuvant systemic therapy.[12]

Although there is great enthusiasm about genomic grading, additional independent validation of the above genomic grade signatures is still required, and their discriminatory power in ER-negative disease has been shown to be rather limited.[34,126] Therefore, despite the issues related to reproducibility of Nottingham histological grade, histological grade

remains a useful and less expensive tool to help tailor the therapy of breast cancer patients. These results confirm the importance of histological grading and should stimulate pathologists to improve intra-observer and interobserver reproducibility.

Response to Therapy: More Signatures!

Despite intense research in the last decades, only 3 biomarkers are mandatory to be tested for adequate breast cancer treatment: ER, PR, and HER2. Molecular studies have confirmed the extreme importance of these biomarkers, and several lines of evidence suggest that breast cancer comprises at least 2 distinct diseases (ER-positive and ER-negative). Investigators aiming to identify new biomarkers and therapeutic targets and/or evaluate response to therapy should take into account the distinctive molecular features of the different subtypes or risk underestimating or overestimating their performance.[26-28] Biological pathways might be important in one particular molecular subtype, such as *TP53* mutation in ER-positive chemotherapy-resistant tumors,[28] whereas the impact of *TP53* mutation in ER-negative disease may differ.

Tamoxifen is one of the few successful examples of targeted therapy in breast cancer. It is the most frequently prescribed endocrine therapy for women with ER-positive breast cancer in both early-stage and advanced disease. However, a substantial proportion of patients relapse because of intrinsic or acquired resistance to tamoxifen. The identification of these patients has been a major goal in breast cancer research. Many groups have tried to identify reliable predictive factors and signatures by means of gene expression profiling in the context of tamoxifen treatment, such as the Oncotype DX[18] and the 2-gene ratio.[17]

Ma et al[17] analyzed the gene expression profile of 60 early-stage tamoxifen-treated patients and identified 3 genes associated with outcome: *HOXB13*, *IL17BR*, and *CHDH*. High levels of *HOXB13* mRNA and low levels of *IL17BR* mRNA were shown to be associated with recurrence. A simple 2-gene ratio *HOXB13:IL17BR* was then proposed as a novel biomarker to predict recurrence in tamoxifen-treated patients. This group developed a qRT-PCR assay and achieved comparable results. Although initial validation studies failed to demonstrate the predictive power of this 2-gene signature,[127] subsequent studies validated the results in 2 different cohorts and showed that the ratio has prognostic power and is predictive of tamoxifen response.[128,129] However, this ratio is only predictive in patients with node-negative, ER-positive cancers,[128,129] which explains its lack of predictive value in the study by Reid et al.[127] Interestingly, further analyses have suggested that the *HOXB13:IL17BR* ratio may have complementary prognostic value when combined with the 5-gene molecular grade index described above.[16]

By analyzing a large series of ER-positive breast cancer patients treated with tamoxifen in the adjuvant setting, Loi et al[15] developed a signature comprising 181 genes that were derived from 13 gene clusters. The classifier was subsequently validated in 3 independent series; however, it was not accurate when applied to a data set of women who had received tamoxifen in the advanced setting. The authors concluded that their signature was mainly prognostic, arguably reflecting the aggressiveness of the tumors. In fact, determining whether a biomarker predicts prognosis or response to therapy is not a simple task. Samples from an advanced breast cancer trial in which patients were randomized to receive tamoxifen or placebo would be required to accurately define the prognostic or predictive nature of this 181-gene signature.[15]

Pre operative and postoperative samples from patients treated with neoadjuvant systemic therapy have recently been subjected to microarray-based expression profiling[11,130-133] and comparative genomic hybridization[134] analyses. Analysis of matched samples provides an excellent opportunity to identify genes whose expression levels change attributable to the systemic therapy.[26] Furthermore, expression profiling analyses of samples from patients achieving pathologic complete response, the best although not ideal surrogate for overall survival in the neoadjuvant setting,[135] can be compared with the analyses of those who did not respond, and predictive signatures for specific drugs or regimens can be devised. In fact, studies based on the above rationale have led to the development of a multigene signature predictive of pathologic complete response to sequential paclitaxel and anthracycline therapy.[11,130] Similar studies using microarray-based or RT-PCR techniques have contributed to our understanding of the mechanisms of response to chemotherapy.[131-133,136-138] Although substantial validation is still needed before these predictors can be incorporated in the clinic, one could argue that in the future, gene expression profiles may help guide the selection of the most appropriate drugs to the correct patients and avoid toxic and ineffective chemotherapy.

Tumor Microenvironment and Stromal Signatures

Apart from the wound-response signature, most of the analyses described above were focused on the epithelial cancer cells. The studies were performed using whole tissue samples, and the samples with insufficient tumor epithelial cell content were generally excluded. Thus, the specific contribution of epithelial and stromal cells to these tumor classifiers is unknown. However, Cleator et al[139] clearly demonstrated that varying proportions of stromal components may influence the accuracy of gene predictors. Lately, attention has been given to the tumor microenvironment and the epithelial-mesenchymal interactions in breast cancer.[8,10,24,140-144] Comprehensive gene expression profiling of each cell type has demonstrated that at the transcriptome level, changes occur in epithelial, myoepithelial, and stromal cells that are already evident at the in situ carcinoma stage.[140] Given the concepts described here, it is perhaps not surprising that prognostic signatures derived from the profiles of stromal cells have been developed and that many genes of the current predictors are predominantly expressed by stromal cells.

By comparing the transcriptome of 2 fibroblastic neoplasia, desmoid-type fibromatosis (DTF) and solitary fibrous tumor, West et al[24] successfully stratified breast cancers into at least 2 prognostically significant subgroups that were independent of traditional clinicopathological factors, suggesting there might be distinct fibroblastic reaction patterns in breast cancer. The cluster, which was enriched for DTF genes, had a significantly better prognosis. Further analyses have confirmed the finding of a good-outcome DTF subgroup in different data sets.[141] Interestingly, at the protein level, breast cancers with strong expression of 1 single gene (*SPARC*) of the DTF cluster in the stromal cells showed a trend for increased survival.[141] However, epithelial cells might express stromal markers, and a study demonstrated that SPARC expression in epithelial cells in a large consecutive series of breast cancers was associated with poor prognosis.[145]

More recently, Finak et al[10] isolated tumor stroma and matched normal stroma from breast tumors and derived a 26-gene signature strongly associated with outcome called the stroma-derived prognostic predictor (SDPP). SDPP predicted prognosis with greater accuracy than the 70-gene signature, stratified several published whole-tumor derived data sets into clinically meaningful subgroups, and was independent of grade, age, lymph node involvement, chemotherapy, hormonal therapy, and ER and HER2 status. The poor-outcome cluster showed markers of hypoxic and angiogenic responses, whereas the good-outcome cluster overexpressed immune-response genes.

Using a different approach, Bergamaschi et al[8] grossly dissected 28 breast cancers and by an unsupervised hierarchical clustering method classified the samples into 4 different groups according to the expression of 278 selected extracellular matrix-related genes. Interestingly, in the validation set, the authors found a strong association between one of the groups (group 1) and the basal-like subgroup previously defined. This group 1 was characterized by high histological grade, lack of ER expression, *TP53* mutations, and poor outcome. Furthermore, 3 of the 278 genes were shown to be prognostically relevant: tumors harboring high expression levels of *MARCO* and low levels of *PUNC* and *SPARC* displayed a significantly worse outcome.[8]

The expression levels of 6 genes pertaining to the extracellular matrix gene cluster (*FN1*, *LOX*, *SPARC*, *TIMP3*, *TNC*, *COL1A1*), which was previously shown to be associated with outcome of breast cancer patients, was analyzed by means of qRT-PCR in a series of 1286 breast cancer patients.[146] It was observed that in a cohort of 680 lymph node–negative patients who did not receive any systemic adjuvant therapy, high levels of *FN1*, *LOX*, and *SPARC* were associated with shorter metastasis-free survival, whereas high levels of *TIMP3* were associated with prolonged metastasis-free survival. *TNC* expression levels were not associated with prognosis in these untreated patients; however, high *TNC* expression was associated with a shorter metastasis-free survival after adjuvant tamoxifen treatment in 139 patients and a shorter progression-free survival in 240 patients whose relapse was treated with first-line tamoxifen monotherapy. These tantalizing results warrant further studies to determine the prognostic and predictive power of these genes. However, one could entertain the hypothesis of testing these genes as potential therapeutic targets of breast cancer.

Signatures Everywhere: Which One Should We Trust?

Apart from the technical issues related to microarray reproducibility and bioinformatic analysis, one issue that has hampered the translation of microarray signatures from interesting scientific observations to clinical tools is the fact that despite the similar prognostic impact of many signatures, the overlap between the gene lists is negligible.[147] For example, only a 3-gene overlap was identified between the 76-gene and 70-gene signatures.[23] As a consequence, both pathologists and clinicians face the challenge of which signature should be used and in what context.

Questions such as whether a single signature would outperform the others and whether combining different signatures would improve the accuracy have been recurrently raised.

Fan and colleagues[148] were the first group undertaking a study to address the concordance between different signatures. These investigators applied multiple classifiers (70-gene signature, 21-gene recurrence score, 2-gene ratio, wound-response signature, and intrinsic gene list) to the same cohort of breast cancer patients and concluded that most of these classifiers showed significant agreement in their outcome prediction for the individual samples. Disappointingly, a model combining different signatures did not perform better than did each one separately. Despite the disparity in their gene lists and the fact that the signatures were developed using different approaches, platforms, and statistical tools, the signatures reflect a common set of phenotypic traits and biological pathways, which are represented in the diverse gene sets. Paraphrasing Massagué,[149] these prognostic signatures "may be regarded as different pictures of the same beast", page 296.

More recently, extensive meta-analyses incorporating publicly available data from several institutions[34,126,147,150,151] have validated the observations made by Fan et al.[148] Wirapati et al[34] confirmed not only that different prognostic signatures identify similar groups of breast cancer patients but also that the assignment of patients into good or poor prognosis is largely dependent on the expression levels of genes pertaining to the "proliferation cluster." Another conclusion drawn from these analyses is that the prognostic power of most classifiers is limited to the ER-positive/HER2-negative subgroup.[34,126] It has been demonstrated that the proliferation cluster is the strongest parameter predicting clinical outcome only in this subtype.[126] Hence, the signatures are essentially detecting as low risk the low-proliferation subset of ER-positive/HER2-negative tumors.[34,126] In fact, Fan et al[148] had already shown that within the luminal A subtype defined by the intrinsic gene list[4,5] there was variability in the outcome predicted by the 4 signatures compared in the study, whereas almost all basal-like, HER2, and luminal B tumors were classified as having a poor outcome. These studies have also demonstrated that some signatures performed even better when only the proliferation-related genes were used to predict prognosis.[34] Thus, some genes might be only introducing noise into the microarray data and may be detrimental to the signatures' performance.

Another recent study compared 13 risk prediction methods, including the expression of a single proliferation gene (*AURKA*), in a training and validation set.[151] Complex signatures consistently failed to outperform the single gene model in the validation set. Therefore, the dramatic increase in variables being tested in parallel by microarrays may not necessarily lead to better predictions. These results challenge once again the approaches used to analyze data generated with high-throughput technologies for the development and validation of breast cancer prognostic and predictive factors. New and more reliable bioinformatic approaches are urgently needed that can extract meaningful statistical and biological information from the large amount of high-dimensional data generated by microarrays in studies characterized by many more interconnected variables (thousands of genes) than samples analyzed (up to hundreds).

Desmedt et al[126] demonstrated that the biological processes that determine the outcome of untreated breast cancer patients vary according to the immunoprofile of the cancers. In ER-positive/HER2-negative cancers, the strongest predictors of outcome are age, size, lymph node metastasis, proliferation as defined by Aurora A expression, and angiogenesis as defined by vascular endothelial growth factor (VEGF) expression, whereas in ER-negative/HER2-negative cancers only immune response as defined by signal transducer and activator of transcription 1 (STAT1) expression is significant (grade, size, and lymph node metastasis are not significant). In HER2 positive cancers, lymph node metastasis, angiogenesis as defined by VEGF expression, tumor invasion as defined by plasminogen activator urokinase expression, and immune response as defined by STAT1 expression are significantly associated to outcome.[126] These results provide even further evidence that the heterogeneity of breast cancer has not only biological implications (ie, breast cancer encompasses multiple diseases that are governed by the expression of different genes and activation/inactivation of distinct pathways) but also clinical relevance.

Estrogen Receptor-negative Disease: Is There a Marker of Good Outcome?

As emphasized above, given the dependence on the expression of genes related to proliferation, the prognostic value of current gene signatures is limited in ER-negative tumors, which usually display a high proliferation rate.[34,126] Thus, to predict prognosis in ER-negative disease, classifiers that rely in other gene modules are needed. One of the stromal signatures described above[10] may have the ability to stratify ER-negative tumors into good and poor prognosis groups. This is in part attributed to enrichment for immune-related genes in the good outcome cluster. Multiple studies have suggested that a lymphocyte-associated gene cluster may identify a group of patients with ER-negative breast cancers who have a good outcome.[21,126,152-155] A recent meta-analysis integrating data of more than 2100 breast cancer patients identified an immune response module as significantly associated with prognosis of ER-negative/ HER2-negative tumors, whereas in HER2-positive tumors an invasion module was also shown to add prognostic information.[126]

Teschendorff et al[155] first identified a 7-gene immune response module that was associated with good outcome in a cohort of 186 ER-negative breast cancer patients. Surprisingly, this was not related to either the extent of lymphocytic infiltration, medullary carcinomas, or *BRCA1* mutation status. Recently, based on these 7 genes and using a complex mathematical algorithm, the same group generated a classifier and validated it in 6 independent ER-negative cohorts.[21] The negative predictive value was >85%; hence, the authors argued, this classifier could be used to identify high-grade ER-negative patients who may benefit from less aggressive systemic therapy.

Conclusion

High throughput molecular techniques have led to a paradigm shift in the way breast cancer is perceived and how breast cancer research is carried out. Although these technologies have proven incredibly useful in basic and translational research studies, their clinical applications are still limited. In this era of tailored therapies, the development of taxonomies that are more predictive is of paramount importance. It is anticipated that by combining morphological, immunohistochemical, and molecular techniques, a more biologically and clinically meaningful classification of breast cancer will emerge. Despite the hype of microarrays and their undeniable contribution to our understanding of breast cancer, important challenges must be met before this methodology is implemented in the clinical setting. The only prospective study in which a microarray-based signature (ie, 70-gene signature) was used to define which patients should receive chemotherapy (ie, MicroarRAy PrognoSTics in Breast CancER [RASTER] study) led to somewhat disconcerting results. First, of 585 eligible patients, 31.6% of patients (n = 158) were excluded because of sampling failure (n = 128) and incorrect procedure (n = 30). Second, it is, therefore, perhaps ironic that although the 70-gene signature was devised to increase the number of patients who could be safely spared the cytotoxic effects of chemotherapy, adjuvant chemotherapy was more often administered to patients whose therapy was guided by the results of the prognostic signature than to patients who were managed solely according to the Dutch CBO clinicopathological guidelines.[92]

With the pace of technology development and the introduction of next-generation sequencers,[156-158] which can provide more accurate, genome-wide, and unbiased, quantitative, and qualitative data on RNA and DNA, it is unlikely that microarrays themselves will ever become one of the main tools for making decisions about treatment for breast cancer patients. Furthermore, the compelling evidence in support of histological grade and proliferation as robust prognostic factors in ER-positive/HER2-negative breast cancers should motivate pathologists to improve the rates of intra-observer and interobserver agreement. Pathologists should not forget that, as emphasized by Rosai,[159] microscopy remains a cornerstone of our specialty; we would go one step further and suggest that histopathology remains a cornerstone of breast cancer patient decision making. However, in this era of tailored therapies and individualized medicine, if we are ever to guide the oncologist's hands as we have been guiding the surgeon's hands, as a community we should strive to play a central role in the implementation and validation of novel molecular techniques for the diagnosis and management of cancer patients.

Acknowledgement

We acknowledge NHS funding to the NIHR Biomedical Research Centre.

References

1. Reis-Filho JS, Simpson PT, Gale T, et al. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. *Pathol Res Pract.* 2005;201: 713-725.
2. Simpson PT, Reis-Filho JS, Gale T, et al. Molecular evolution of breast cancer. *J Pathol.* 2005;205:248-254.
3. Nahta R, Esteva FJ. Trastuzumab: triumphs and tribulations. *Oncogene* 2007;26:3637-3643.
4. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747-752.

5. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001;98:10869-10874.
6. Hu Z, Fan C, Oh DS, et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics*. 2006;7:96.
7. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*. 2003;100:8418-8423.
8. Bergamaschi A, Tagliabue E, Sorlie T, et al. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. *J Pathol*. 2008;214:357-367.
9. 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:E7.
10. Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med*. 2008;14:518-527.
11. Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol*. 2006;24:4236-4244.
12. Ivshina AV, George J, Senko O, et al. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res*. 2006;66:10292-10301.
13. Liu J, Campen A, Huang S, et al. Identification of a gene signature in cell cycle pathway for breast cancer prognosis using gene expression profiling data. *BMC Med Genom*. 2008;1:39.
14. 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;356:217-226.
15. Loi S, Haibe-Kains B, Desmedt C, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics*. 2008;9:239.
16. Ma XJ, Salunga R, Dahiya S, et al. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. *Clin Cancer Res*. 2008;14:2601-2608.
17. 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;5:607-616.
18. 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;351:2817-2826.
19. Ring BZ, Seitz RS, Beck R, et al. Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer. *J Clin Oncol*. 2006;24:3039-3047.
20. 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;98:262-272.
21. Teschendorff AE, Caldas C. A robust classifier of high predictive value to identify good prognosis patients in ER negative breast cancer. *Breast Cancer Res*. 2008;10:R73.
22. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415:530-536.
23. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet*. 2005;365:671-679.
24. West RB, Nuyten DS, Subramanian S, et al. Determination of stromal signatures in breast carcinoma. *PLoS Biol*. 2005;3:e187.
25. He YD, Friend SH. Microarrays—the 21st century divining rod? *Nat Med*. 2001;7:658-659.
26. Andre F, Mazouni C, Hortobagyi GN, et al. DNA arrays as predictors of efficacy of adjuvant/neoadjuvant chemotherapy in breast cancer patients: current data and issues on study design. *Biochim Biophys Acta*. 2006;1766:197-204.
27. Andre F, Pusztai L. Molecular classification of breast cancer: implications for selection of adjuvant chemotherapy. *Nat Clin Pract Oncol*. 2006;3:621-632.
28. Tordai A, Wang J, Andre F, et al. Evaluation of biological pathways involved in chemotherapy response in breast cancer. *Breast Cancer Res*. 2008;10:R37.
29. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol*. 2007;25:5287-5312.
30. Dunkler D, Michiels S, Schemper M. Gene expression profiling: does it add predictive accuracy to clinical characteristics in cancer prognosis? *Eur J Cancer*. 2007;43:745-751.
31. Lu X, Lu X, Wang ZC, et al. Predicting features of breast cancer with gene expression patterns. *Breast Cancer Res Treat*. 2008;108:191-201.
32. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet*. 2005;365:488-492.
33. Michiels S, Koscielny S, Hill C. Interpretation of microarray data in cancer. *Br J Cancer*. 2007;96:1155-1158.
34. Wirapati P, Sotiriou C, Kunkel S, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res*. 2008;10:R65.
35. Brenton JD, Carey LA, Ahmed AA, et al. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? *J Clin Oncol*. 2005;23:7350-7360.
36. Reis-Filho JS, Westbury C, Pierga JY. The impact of expression profiling on prognostic and predictive testing in breast cancer. *J Clin Pathol*. 2006;59:225-231.

37. Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res.* 2005;11:5678-5685.
38. Geyer CE, Forster J, Lindquist D, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *N Engl J Med.* 2006;355:2733-2743.
39. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med.* 2006;354: 809-820.
40. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med.* 2005;353:1659-1672.
41. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med.* 2005;353:1673-1684.
42. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med.* 2001;344:783-792.
43. Abd El-Rehim DM, Pinder SE, Paish CE, et al. Expression of luminal and basal cytokeratins in human breast carcinoma. *J Pathol*. 2004;203:661-671.
44. Gusterson BA, Ross DT, Heath VJ, et al. Basal cytokeratins and their relationship to the cellular origin and functional classification of breast cancer. *Breast Cancer Res.* 2005;7:143-148.
45. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res.* 2004;10:5367-5374.
46. Rakha EA, Putti TC, Abd El-Rehim DM, et al. Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *J Pathol.* 2006;208:495-506.
47. Savage K, Lambros MB, Robertson D, et al. Caveolin 1 is overexpressed and amplified in a subset of basal-like and metaplastic breast carcinomas: a morphologic, ultrastructural, immunohistochemical, and in situ hybridization analysis. *Clin Cancer Res.* 2007;13:90-101.
48. Savage K, Leung S, Todd SK, et al. Distribution and significance of caveolin 2 expression in normal breast and invasive breast cancer: an immunofluorescence and immunohistochemical analysis. *Breast Cancer Res Treat.* 2008;110:245-256.
49. Turner NC, Reis-Filho JS. Basal-like breast cancer and the BRCA1 phenotype. *Oncogene* 2006;25:5846-5853.
50. van de Rijn M, Perou CM, Tibshirani R, et al. Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. *Am J Pathol.* 2002;161: 1991-1996.
51. Parry S, Savage K, Marchio C, et al. Nestin is expressed in basal-like and triple-negative breast cancers. *J Clin Pathol.* 2008;61:1045-1050.
52. Gilbert JA, Goetz MP, Reynolds CA, et al. Molecular analysis of metaplastic breast carcinoma: high EGFR copy number via aneusomy. *Mol Cancer Ther.* 2008;7: 944-951.
53. Fulford LG, Easton DF, Reis-Filho JS, et al. Specific morphological features predictive for the basal phenotype in grade 3 invasive ductal carcinoma of breast. *Histopathology*. 2006;49:22-34.
54. Livasy CA, Karaca G, Nanda R, et al. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Mod Pathol.* 2006;19:264-271.
55. Reis-Filho JS, Milanezi F, Steele D, et al. Metaplastic breast carcinomas are basal-like tumours. *Histopathology*. 2006;49:10-21.
56. Turner NC, Reis-Filho JS, Russell AM, et al. BRCA1 dysfunction in sporadic basal-like breast cancer. *Oncogene* 2007;26:2126-2132.
57. Liu X, Holstege H, van der Gulden H, et al. Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer. *Proc Natl Acad Sci U S A.* 2007;104:12111-12116.
58. McCarthy A, Savage K, Gabriel A, et al. A mouse model of basal-like breast carcinoma with metaplastic elements. *J Pathol.* 2007;211:389-398.
59. Dairkee SH, Mayall BH, Smith HS, et al. Monoclonal marker that predicts early recurrence of breast cancer. *Lancet.* 1987;1:514.
60. Gould VE, Koukoulis GK, Jansson DS, et al. Coexpression patterns of vimentin and glial filament protein with cytokeratins in the normal, hyperplastic, and neoplastic breast. *Am J Pathol.* 1990;137:1143-1155.
61. Malzahn K, Mitze M, Thoenes M, et al. Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas. *Virchows Arch* 1998;433:119-129.
62. Murad TM, Scharpelli DG. The ultrastructure of medullary and scirrhous mammary duct carcinoma. *Am J Pathol.* 1967;50:335-360.
63. Santini D, Ceccarelli C, Taffurelli M, et al. Differentiation pathways in primary invasive breast carcinoma as suggested by intermediate filament and biopathological marker expression. *J Pathol.* 1996;179:386-391.
64. Banerjee S, Reis-Filho JS, Ashley S, et al. Basal-like breast carcinomas: clinical outcome and response to chemotherapy. *J Clin Pathol.* 2006;59:729-735.
65. Rakha EA, El-Sayed ME, Green AR, et al. Prognostic markers in triple-negative breast cancer. *Cancer* 2007; 109:25-32.
66. Rakha EA, Tan DS, Foulkes WD, et al. Are triple-negative tumours and basal-like breast cancer synonymous? *Breast Cancer Res.* 2007;9:404; author reply 405.
67. Bertucci F, Finetti P, Cervera N, et al. How basal are triple-negative breast cancers? *Int J Cancer*. 2008;123: 236-240.
68. Cheang MC, Voduc D, Bajdik C, et al. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin Cancer Res.* 2008;14:1368-1376.

69. Rakha EA, Reis-Filho JS, Ellis IO. Basal-like breast cancer: a critical review. *J Clin Oncol*. 2008;26:2568-2581.
70. Rakha EA, Reis-Filho JS, Ellis IO. Impact of basal-like breast carcinoma determination for a more specific therapy. *Pathobiology*. 2008;75:95-103.
71. Calza S, Hall P, Auer G, et al. Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. *Breast Cancer Res.* 2006;8:R34.
72. Jumppanen M, Gruvberger-Saal S, Kauraniemi P, et al. Basal-like phenotype is not associated with patient survival in estrogen-receptor-negative breast cancers. *Breast Cancer Res.* 2007;9:R16.
73. Tan DS, Marchió C, Jones RL, Savage K, Smith IE, Dowsett M, Reis-Filho JS. Triple-negative breast cancer: molecular profiling and prognostic impact in adjuvant anthracycline-treated patients. *Breast Cancer Res Treat.* 2008;111:27-44.
74. Mullins M, Perreard L, Quackenbush JF, et al. Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues. *Clin Chem.* 2007;53:1273-1279.
75. Parker J, Mullins M, Cheang MC, et al. A supervised risk predictor of breast cancer based on biological subtypes. ASCO Meeting Abstract May 20 2008: 11008.
76. Perreard L, Fan C, Quackenbush JF, et al. Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. *Breast Cancer Res.* 2006;8:R23.
77. Farmer P, Bonnefoi H, Becette V, et al. Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene*. 2005;24:4660-4671.
78. Herschkowitz JI, Simin K, Weigman VJ, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol.* 2007;8:R76.
79. Weigelt B, Horlings H, Kreike B, et al. Refinement of breast cancer classification by molecular characterization of histological special types. *J Pathol.* 2008;216:141-150.
80. Marchio C, Iravani M, Natrajan R, et al. Genomic and immunophenotypical characterization of pure micropapillary carcinomas of the breast. *J Pathol.* 2008;215:398-410.
81. Abdel-Fatah TM, Powe DG, Hodi Z, et al. High frequency of coexistence of columnar cell lesions, lobular neoplasia, and low grade ductal carcinoma in situ with invasive tubular carcinoma and invasive lobular carcinoma. *Am J Surg Pathol*. 2007;31:417-426.
82. Abdel-Fatah TM, Powe DG, Hodi Z, et al. Morphologic and molecular evolutionary pathways of low nuclear grade invasive breast cancers and their putative precursor lesions: further evidence to support the concept of low nuclear grade breast neoplasia family. *Am J Surg Pathol*. 2008;32:513-523.
83. Weigelt B, Kreike B, Reis-Filho JS. Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis. *Breast Cancer Res Treat.* 2008 Sep 25. [Epub ahead of print]
84. Tognon C, Knezevich SR, Huntsman D, et al. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer Cell.* 2002;2: 367-376.
85. Diallo R, Schaefer KL, Bankfalvi A, et al. Secretory carcinoma of the breast: a distinct variant of invasive ductal carcinoma assessed by comparative genomic hybridization and immunohistochemistry. *Hum Pathol.* 2003;34:1299-1305.
86. Diallo R, Tognon C, Knezevich SR, et al. Secretory carcinoma of the breast: a genetically defined carcinoma entity. *Verh Dtsch Ges Pathol.* 2003;87:193-203.
87. Makretsov N, He M, Hayes M, et al. A fluorescence in situ hybridization study of ETV6-NTRK3 fusion gene in secretory breast carcinoma. *Genes Chromosomes Cancer*. 2004;40:152-157.
88. Reis-Filho JS, Natrajan R, Vatcheva R, et al. Is acinic cell carcinoma a variant of secretory carcinoma? A FISH study using ETV6 "split apart" probes. *Histopathology*. 2008;52:840-846.
89. Galea MH, Blamey RW, Elston CE, et al. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat.* 1992;22:207-219.
90. Goldhirsch A, Wood WC, Gelber RD, et al. Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. *Ann Oncol*. 2007;18:1133-1144.
91. Eifel P, Axelson JA, Costa J, et al. National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000. *J Natl Cancer Inst*. 2001;93:979-989.
92. Bueno-de-Mesquita JM, van Harten WH, Retel VP, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). *Lancet Oncol.* 2007;8:1079-1087.
93. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst*. 2006;98: 1183-1192.
94. Mook S, Schmidt MK, Viale G, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study. *Breast Cancer Res Treat.* 2008 Jul 27. [Epub ahead of print]
95. 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;347:1999-2009.
96. Cardoso F, Van't Veer L, Rutgers E, et al. Clinical application of the 70-gene profile: the MINDACT trial. *J Clin Oncol*. 2008;26:729-735.
97. Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative

breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res.* 2007;13: 3207-3214.

98. 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;24:1665-1671.
99. Kreike B, van Kouwenhove M, Horlings H, et al. Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. *Breast Cancer Res.* 2007;9:R65.
100. 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 U S A*. 2005;102:3738-3743.
101. Al-Hajj M, Wicha MS, Benito-Hernandez A, et al. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A*. 2003;100:3983-3988.
102. Ioannidis JP. Microarrays and molecular research: noise discovery? *Lancet*. 2005;365:454-455.
103. Eden P, Ritz C, Rose C, et al. “Good old” clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. *Eur J Cancer*. 2004;40:1837-1841.
104. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc Natl Acad Sci U S A*. 2006;103: 5923-5928.
105. Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol*. 2005;23:7332-7341.
106. Simon R, Radmacher MD, Dobbin K, et al. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst.* 2003;95:14-18.
107. Shi L, Reid LH, Jones WD, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol.* 2006;24:1151-1161.
108. Penland SK, Keku TO, Torrice C, et al. RNA expression analysis of formalin-fixed paraffin-embedded tumors. *Lab Invest.* 2007;87:383-391.
109. Ravo M, Mutarelli M, Ferraro L, et al. Quantitative expression profiling of highly degraded RNA from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays. *Lab Invest.* 2008;88: 430-440.
110. 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:R25.
111. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol.* 2006;24:3726-3734.
112. Goldstein LJ, Gray R, Badve S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. *J Clin Oncol.* 2008;26:4063-4071.
113. Flanagan MB, Dabbs DJ, Brufsky AM, et al. Histopathologic variables predict Oncotype DXtrade mark Recurrence Score. *Mod Pathol.* 2008;21:1255-1261.
114. Paik S. Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with tamoxifen. *Oncologist*. 2007;12: 631-635.
115. Ross JS, Hatzis C, Symmans WF, et al. Commercialized multigene predictors of clinical outcome for breast cancer. *Oncologist*. 2008;13:477-493.
116. Wolf I, Ben-Baruch N, Shapira-Frommer R, et al. Association between standard clinical and pathologic characteristics and the 21-gene recurrence score in breast cancer patients: a population-based study. *Cancer* 2008;112:731-736.
117. Buerger H, Mommers EC, Littmann R, et al. Ductal invasive G2 and G3 carcinomas of the breast are the end stages of at least two different lines of genetic evolution. *J Pathol.* 2001;194:165-170.
118. Buerger H, Otterbach F, Simon R, et al. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol.* 1999;189:521-526.
119. Roylance R, Gorman P, Harris W, et al. Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. *Cancer Res.* 1999;59:1433-1436.
120. Ma XJ, Salunga R, Tuggle JT, et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A*. 2003;100:5974-5979.
121. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer, I: the value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;19:403-410.
122. Robbins P, Pinder S, de Klerk N, et al. Histological grading of breast carcinomas: a study of interobserver agreement. *Hum Pathol.* 1995;26:873-879.
123. Loi S, Haibe-Kains B, Desmedt C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol.* 2007;25:1239-1246.
124. Desmedt C, Sotiriou C. Proliferation: the most prominent predictor of clinical outcome in breast cancer. *Cell Cycle*. 2006;5:2198-2202.
125. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev.* 2007;7:545-553.
126. Desmedt C, Haibe-Kains B, Wirapati P, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clin Cancer Res.* 2008;14:5158-5165.
127. Reid JF, Lusa L, De Cecco L, et al. Limits of predictive models using microarray data for breast cancer clinical treatment outcome. *J Natl Cancer Inst.* 2005;97:927-930.

128. Goetz MP, Suman VJ, Ingle JN, et al. A two-gene expression ratio of homeobox 13 and interleukin-17B receptor for prediction of recurrence and survival in women receiving adjuvant tamoxifen. *Clin Cancer Res.* 2006;12:2080-2087.
129. 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;24:4611-4619.
130. 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;22:2284-2293.
131. Chang JC, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet.* 2003;362:362-369.
132. 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;11:7434-7443.
133. 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;23:422-431.
134. Pierga JY, Reis-Filho JS, Cleator SJ, et al. Microarray-based comparative genomic hybridisation of breast cancer patients receiving neoadjuvant chemotherapy. *Br J Cancer.* 2007;96:341-351.
135. Kuerer HM, Newman LA, Smith TL, et al. Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. *J Clin Oncol.* 1999;17:460-469.
136. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J Clin Oncol.* 2005;23:7265-7277.
137. Lee JK, Havaleshko DM, Cho H, et al. A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. *Proc Natl Acad Sci U S A.* 2007;104:13086-13091.
138. Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. *Nat Med.* 2006;12:1294-1300.
139. Cleator SJ, Powles TJ, Dexter T, et al. The effect of the stromal component of breast tumours on prediction of clinical outcome using gene expression microarray analysis. *Breast Cancer Res.* 2006;8:R32.
140. Allinen M, Beroukhim R, Cai L, et al. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell.* 2004;6:17-32.
141. Beck AH, Espinosa I, Gilks CB, et al. The fibromatosis signature defines a robust stromal response in breast carcinoma. *Lab Invest.* 2008;88:591-601.
142. Casey T, Bond J, Tighe S, et al. Molecular signatures suggest a major role for stromal cells in development of invasive breast cancer. *Breast Cancer Res Treat.* 2008 Mar 29. [Epub ahead of print]
143. Hu M, Yao J, Carroll DK, et al. Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell.* 2008;13:394-406.
144. Weigelt B, Bissell MJ. Unraveling the microenvironmental influences on the normal mammary gland and breast cancer. *Semin Cancer Biol.* 2008;18:311-21.
145. Jones C, Mackay A, Grigoriadis A, et al. Expression profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer. *Cancer Res.* 2004;64:3037-3045.
146. Helleman J, Jansen MP, Ruigrok-Ritstier K, et al. Association of an extracellular matrix gene cluster with breast cancer prognosis and endocrine therapy response. *Clin Cancer Res.* 2008;14:5555-5564.
147. Thomassen M, Tan Q, Eiriksdottir F, et al. Comparison of gene sets for expression profiling: prediction of metastasis from low-malignant breast cancer. *Clin Cancer Res.* 2007;13:5355-5360.
148. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med.* 2006;355:560-569.
149. Massagué J. Sorting out breast-cancer gene signatures. *N Engl J Med.* 2007;356:294-297.
150. Haibe-Kains B, Desmedt C, Piette F, et al. Comparison of prognostic gene expression signatures for breast cancer. *BMC Genom.* 2008;9:394.
151. Haibe-Kains B, Desmedt C, Sotiriou C, et al. A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? *Bioinformatic.* 2008;24:2200-2208.
152. Alexe G, Dalgin GS, Scanfeld D, et al. High expression of lymphocyte-associated genes in node-negative HER2+ breast cancers correlates with lower recurrence rates. *Cancer Res.* 2007;67:10669-10676.
153. Calabro A, Beissbarth T, Kuner R, et al. Effects of infiltrating lymphocytes and estrogen receptor on gene expression and prognosis in breast cancer. *Breast Cancer Res Treat.* 2008 Jul 1. [Epub ahead of print]
154. Schmidt M, Bohm D, von Torne C, et al. The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res.* 2008;68:5405-5413.
155. Teschendorff AE, Miremadi A, Pinder SE, et al. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. *Genome Biol.* 2007;8:R157.
156. Kahvejian A, Quackenbush J, Thompson JF. What would you do if you could sequence everything? *Nat Biotechnol.* 2008;26:1125-1133.
157. Ledford H. The death of microarrays? *Nature* 2008; 455:847.
158. Shendure J. The beginning of the end for microarrays? *Nat Methods.* 2008;5:585-587.
159. Rosai J. Why microscopy will remain a cornerstone of surgical pathology. *Lab Invest.* 2007;87:403-408.