

Proteomics of Breast Cancer

PRINCIPLES AND POTENTIAL CLINICAL APPLICATIONS*

François Bertucci‡§¶, Daniel Birnbaum‡, and Anthony Goncalves§¶**

Progresses in screening, early diagnosis, prediction of aggressiveness and of therapeutic response or toxicity, and identification of new targets for therapeutic will improve survival of breast cancer. These progresses will likely be accelerated by the new proteomic techniques. In this review, we describe the different techniques currently applied to clinical samples of breast cancer and the most important results obtained with the two most popular proteomic approaches in translational research (tissue microarrays and SELDI-TOF). *Molecular & Cellular Proteomics* 5:1772–1786, 2006.

Worldwide, breast cancer is one of the most frequent and deadly cancers. Although the survival of patients has increased over the last decades in relation with screening programs and postoperative adjuvant systemic therapies (hormone therapy and chemotherapy), many patients die from metastatic relapse. Breast cancer is a complex disease. Accumulation of numerous and often unknown molecular alterations causes cell proliferation, genetic instability, and acquisition of an increasingly invasive and resistant phenotype. The combinatorial origin and the heterogeneity of malignant cells and the variability of the host background create molecularly distinct subgroups of tumors endowed with different phenotypes and clinical outcomes. This heterogeneity is only partially apprehended by the clinicopathological parameters currently used by clinicians, hampering the selection of the most appropriate diagnostic or therapeutic strategy for each case. In addition, current cytotoxic drugs do not differentiate between cancerous and normal cells, causing sometimes disastrous adverse effects. Major efforts aim at characterizing the heterogeneity of the disease by defining more reliable diagnostic/prognostic factors and at developing molecular therapies selectively targeting the tumor cells.

For some decades, the study of molecular alterations has successfully elucidated some mechanisms of mammary oncogenesis and identified key genes such as *ERBB2*, *TP53*, *CCND1*, *BRCA1*, and *BRCA2*. It has also allowed consider-

able therapeutic progress by targeting hormonal receptors and ERBB2/HER2 receptor. Today high throughput molecular typing provides an unprecedented opportunity to tackle the combinatory aspect and the complexity of breast cancer. A combination of markers (molecular signature) will probably be more sensitive and specific than a single molecular marker to reflect the actual heterogeneity of disease; more reliable for screening, diagnosis, prognosis, and prediction of therapeutic response; and more useful to find new therapeutic targets. The first large scale techniques applied to the cancer field were DNA microarrays for mRNA expression profiling (1). Several studies have demonstrated the clinical potential of this approach, notably by identifying new biologically relevant and prognostic subclasses of breast cancer unidentifiable by conventional means (2–14). With the recent development of similar technologies at the DNA and protein levels, there is no reason to limit profiling to RNA.

Proteomics has the potential to complement and further enlarge the wealth of information generated by genomics in breast cancer for several reasons. mRNA levels do not necessarily correlate with corresponding protein abundance (15–19). Additional complexity is conferred by protein post-translational modifications, including phosphorylations, acetylations, and glycosylations, or protein cleavages (20). These modifications are not detectable at the mRNA level but play significant roles in protein stability, localization, interactions, and functions. Finally proteins represent more accessible and relevant therapeutic targets than nucleic acids.

However, for a long time, protein measurements have been considered more difficult than DNA or RNA analyses. Composed of 21 naturally occurring amino acids, proteins can represent a virtually unlimited variety of molecules that has been estimated from 1 to 20 million in mammalian organisms. Moreover protein abundance in a biological sample may vary from few copies to several millions. Consequently an accurate quantitative estimation, most often from limited material in clinical samples, requires both high sensitivity and dynamic range. Although these requirements are only partially achieved yet, several technical strategies have been developed during the past 10 years. They allow access to the proteome in various biological samples with still preliminary but promising results in the clinical setting. Breast cancer studies have concerned tumor tissue but also biological fluids including serum, plasma, nipple aspirate, or ductal lavage material with different intents, including better understanding of mammary oncogenesis and improvements in screening

From the ‡Centre de Recherche en Cancérologie de Marseille, Département d'Oncologie Moléculaire, Institut Paoli-Calmettes (IPC) and UMR599 INSERM, §Département d'Oncologie Médicale and **Département de Pharmacologie Moléculaire, IPC, and ¶Faculté de Médecine, Université de la Méditerranée, 13009 Marseille, France

Received, May 24, 2006

Published, MCP Papers in Press, May 29, 2006, DOI 10.1074/mcp.R600011-MCP200

and diagnosis as well as prognosis and/or prediction of therapeutic response.

In this review, we will first present the main characteristics of proteomics-based methods that have been applied to date to breast cancer samples. Then we will discuss the preliminary results generated from clinical samples by the most common approaches and the perspectives of protein profiling.

PROTEOMICS METHODS

The proteomic techniques may be roughly distinguished according to the requirement or not for an initial biological foreknowledge. Approaches that require “*a priori*” hypotheses or selections include microarray-based techniques. They are essentially dominated by tissue microarrays (TMAs),¹ which address protein expression of candidate markers on a large number of arrayed tumor tissues. Other microarray-based techniques include antibody arrays and reverse-phase protein arrays, which are promising methods but still in their infancies with respect to clinical applications. Approaches without a *priori* requirement, which do not suppose any previous biological foreknowledge, allow examining and quantifying a large number of initially unknown protein parameters. They are used to identify unique or multiple protein markers, which correlate with a pertinent tumor phenotype. This kind of approach is essentially based on mass spectrometry tools, most often following an initial chemical or physical separation of the analyte.

Tissue Microarrays

Increasing identification of molecular markers with potential clinical significance by high throughput techniques has boosted the development of fast approaches to validation in large series of samples. DNA or protein microarrays have some disadvantages (high cost and requirement of unfixed frozen samples) that make them not adaptable for such validation. Initially described ~20 years ago (21, 22), TMA was modernized in 1998 (23). The technique allows the simultaneous analysis at the DNA (fluorescence *in situ* hybridization), mRNA (*in situ* hybridization), or protein (immunohistochemistry (IHC)) levels of up to 1,000 tumor samples arrayed onto one microscope glass slide. Cylindrical cores as small as 600 μm in diameter are generated from formalin-fixated paraffin-embedded (FFPE) archived tumor blocks. These cores are arrayed in a new paraffin block that is then cut into thin sections (100–200/block) for analysis. All samples from a section may be simultaneously interrogated by a specific

antibody by IHC as well as morphologically analyzed. Immunostaining is scored by the pathologist. In contrast to the other methods described in this review, the molecular information is obtained *in situ* in the context of cell morphology and tissue architecture, and specimens are derived from FFPE archival tissues (although TMAs may also be done from frozen tissues (cryo-TMAs)). Advantages of TMA are multiple as compared with traditional IHC on whole sections. TMA is much less time-consuming, labor-intensive, and economically costly; all samples are tested in similar experimental conditions, leading to more reproducible and rigorous data; and the equipment required to make TMAs is not complicated. Finally TMA helps in the management of tissue archives by allowing preservation of tissue resources for research as well as in the testing and optimization of diagnostic tests.

TMA is so far the most used proteomic technique in oncology with more than 1,500 publications in PubMed in April 2006 with the terms “tissue microarrays” and “cancer.” The high number of samples simultaneously profiled reinforces the statistical significance of results. Analysis of individual molecular markers on large series of samples remains the main application of the technique with different objectives with regard to the spotted samples. TMA has been used for the screening of multiple tumor types (“multitumor TMAs”) (24) for the expression of a protein of interest (25, 26). The largest reported multitumor TMA, constructed at the University of Basel, contained up to 4,700 tumors representing 135 different tumor types (27). Another objective is the analysis of different stages of tumor progression for a given cancer as reported for bladder (28, 29) or prostate cancers (30, 31). If follow-up data are available, TMAs may be used to measure the prognostic significance of a given protein as reported in colorectal (32, 33), renal (34), and bladder (28) cancers. TMAs can also be constructed from cell lines (35) or other experimental material such as xenografts (31). In these applications, the tested individual markers have previously known or suspected clinical relevance or represent novel markers derived from high throughput molecular analyses. More recently, TMAs were used to identify among tens of tested proteins a multiprotein expression signature associated with a given phenotype. The multidimensionality of the dataset imposed the use of complex analytic tools similar to those used for analysis of quantitative DNA microarray data (36), but the task was complicated by the qualitative or semiquantitative nature of the IHC data. Unsupervised hierarchical clustering was first applied and allowed the identification of coherent tumor and protein clusters. Until now, supervised methods have not been frequently used to define multiprotein predictors of a given phenotype. These multiprotein approaches were reported in lymphoma (37), uterine (38), and prostate (39) cancers; melanoma (40); and sarcoma (41).

Several criticisms have been formulated with regard to the use of TMAs in the cancer field. The most recurrent concerned the representativity of a small sample (0.6 mm in diameter) of

¹ The abbreviations used are: TMA, tissue microarray; IHC, immunohistochemistry; FFPE, formalin-fixated paraffin-embedded; NAF, nipple aspirate fluid; MBC, medullary breast carcinoma; IBC, inflammatory breast cancer; NIBC, non-IBC; EBC, early breast cancer; MFS, metastasis-free survival; ER, estrogen receptor; 2D, two-dimensional; EGFR, epidermal growth factor receptor.

a potentially heterogeneous tumor as compared with a classical large section. In fact, several studies (17, 42) have shown excellent concordance for IHC when two or three samples per tumor were carefully selected by the pathologist for TMA construction. Other criticisms include those related to IHC (subjectivity and reproducibility of manual scoring; qualitative or semiquantitative scoring, which may lack clinically relevant information; inconstant availability of antibodies with sufficient quality for analysis of FFPE samples; and lack of standardization of protocols), the prior knowledge of the proteins to be tested that does not readily contribute to the discovery of new markers as well as the need for bioinformatical tools for analysis of multidimensional qualitative data, or the frequent discordance between mRNA and protein expression levels. In the future, a much more widespread use of TMA is expected and will help minimize these limitations. For example, the automated quantitative analysis of TMA has made progress, providing more rapid, reproducible, and objective results (44–47); several software packages are commercially available (48), and such automation will be a key issue in the future. Yet the first results of TMA studies are extremely encouraging, notably in breast cancer. Results are described below.

Protein Microarrays

These approaches tend to recapitulate at the protein level the mRNA expression profiling studies by arraying various protein probes onto specific surfaces and then measuring interactions with specific proteins in complex samples. The most advanced format in this setting is the antibody microarray where the bait proteins are specific antibodies printed on solid surfaces (49). Samples to analyze are either directly labeled with fluorophores or indirectly labeled with a tag subsequently detected by a labeled antibody. A potential problem is that direct protein labeling may disrupt the binding of an analyte. In addition, because all proteins are labeled, high abundance proteins may bind nonspecifically to the bait molecule, generating high background. A reference sample may be co-incubated with a test sample to normalize for variation between spots in capture antibody concentration. The assay is competitive and generates a linear response according to the concentration of the analyte. This two-color comparative fluorescence strategy has been used to compare protein levels between malignant and normal breast tissues from the same patient (50). It used a commercially available antibody microarray (Clontech, Mountain View, CA) targeting 378 different proteins. Several proteins had higher levels in the malignant tissue, which was confirmed by IHC. A similar approach using the Panorama™ Ab Microarray-Cell Signaling array (Sigma) has been applied to adipose tissue of mammary gland from high risk breast cancer patients and revealed the expression of a large number of proteins potentially involved in the regulation of the breast tumor microenvironment (51).

Alternatively labeling can be done after array incubation

with a second specific antibody supposed to bind the same protein as the bait antibody but on a different epitope. This antibody can be linked to a detection system (“sandwich” method) such as fluorophores, resonance light scattering, enhanced chemiluminescence, tyramide signal amplification, and streptavidin systems. This kind of strategy has been applied to biological fluids related to breast cancer samples. Medium from human breast cancer cells in culture was incubated on cytokine antibody array membranes, revealed by biotin-tagged antibodies, horseradish peroxidase-conjugated streptavidin, and the ECL system. Interleukin-8 was associated with ER status and metastatic potential of breast cancer cells (52). Similarly the expression of various cytokines was studied using cytokine-dedicated antibody microarrays in serum, tumor, or fat tissue interstitial fluids from patients (51, 53, 54).

Protein and antibody arrays may also provide information on the post-translational modifications of targeted proteins. For example, secondary antibodies used in sandwich assays may target either total proteins or the active phosphotyrosine forms. This has been applied to breast cancer cell lysates to evaluate the activation of ERBB receptors upon ligand binding and after specific perturbation by receptor inhibitors (55).

Reverse-phase protein array is another promising approach (56). Protein samples are robotically spotted onto specific surfaces and then probed by appropriate antibodies with labeled secondary antibodies and a detection/amplification system. This method, like the TMAs, allows various samples to be analyzed in a single experiment but with a limited throughput of antibodies to test.

A number of reports have demonstrated the utility of protein arrays, but their wide application in cancer research is still limited in part because of the cost of producing antibodies and the limited availability of antibodies with high specificity and affinity for the target. As already mentioned, antibody arrays are generally hypothesis-driven approaches, which require a previous knowledge of a limited number of candidate proteins to evaluate their expression among different samples. New high throughput antibody production and screening methods could allow efficient probing of a large number of uncharacterized proteins. This has been demonstrated with phage display-based strategies where a large library of peptides can be expressed by phages and then secondarily probed with cancer patient serum to identify novel tumor antigen peptides with immunoreactivity associated to malignant disease (57) as well as autoantibodies with potential diagnostic values (58).

Mass Spectrometry-based Approaches

These approaches take advantage of the ability of mass spectrometers to separate peptides or proteins according to their m/z . Mass spectrometer measurements may serve to identify peptides after tryptic digestion of proteins separated

from complex samples or may be applied to more complex samples to generate a protein signature that correlates with a given phenotype.

Gel-based Approaches—Mass spectrometry-based proteomics may include an initial phase of protein separation. The standard is 2D electrophoresis (59). In 2D electrophoresis, proteins are separated according to two independent physicochemical parameters, isoelectric focusing and molecular weight. After successful separation and appropriate gel staining, proteins of interest are isolated from the gel, digested, and subjected to mass spectrometry. Peptide mass mapping by MALDI-TOF MS analysis is now a routine tool for the identification of proteins separated by 2D electrophoresis (60). MALDI-TOF MS calculates the time-of-flight of peptides subjected to analysis, which is correlated to their m/z . These peptide masses are used to search sequence databases and match with a known protein. The efficiency of MS is dependent on the development of comprehensive sequence databases and expressed sequence tag databases. When the peptide mass mapping experiment does not yield an unequivocal match, fragment ion measurement by MS/MS may be done. Commercial availability of tandem mass spectrometers with MALDI sources, such as the Q-TOF and TOF-TOF, also allows the acquisition of more extensive fragmentation for one or more peptides to confirm the peptide mass map results. ESI-MS/MS, an alternative to MALDI in terms of ion production (61), is also very frequently used coupled to triple quadrupole, quadrupole ion trap, and quadrupole TOF to generate peptide fragmentation spectra. Peptide digests may also be analyzed by LC-MS/MS. To improve speed, sensitivity, and reproducibility of conventional 2D electrophoresis, 2D DIGE was developed (62). With this method, protein extracts are covalently labeled using different fluorescent dyes, e.g. cyanine (Cy2, Cy3, or Cy5) dyes or Alexa dye. Typically the test sample is labeled with Cy3, and the reference sample is labeled with Cy5. Equal concentrations of the two differentially labeled protein samples are then mixed and co-separated during the same 2D electrophoresis process. The 2D DIGE pattern is then visualized by scanning the gel at two wavelengths (for Cy3 and Cy5) using a fluorescence imager. A comparison of the two images generated allows the quantification of each spot. 2D DIGE eliminates or drastically reduces gel-to-gel variability associated with standard 2D electrophoresis and improves the accuracy of quantitative protein profiling, allowing its use to be envisioned in a high throughput research environment (63).

A large number of 2D electrophoresis-based approaches have been developed in preclinical models of breast cancer, including cell lines, but relatively few data have been generated from clinical samples and generally from a limited number of samples. Studies have described *in situ* and/or infiltrating ductal carcinoma and compared them with normal tissues or non-malignant tumors (64–71). These approaches have also been applied to biological fluids including serum (72) and

nipple aspirate fluids (NAFs) (73). Although 2D electrophoresis has excellent resolution, the requirement for subsequent protein staining and sample handling limits the exhaustivity, sensitivity, and reproducibility of the approach. Other major drawbacks that limit its clinical application include the requirement of a large amount of starting material and its labor-intensive nature.

Non-gel-based Approaches—Other MS-based techniques use separative techniques that are not dependent on gel electrophoresis. A two-dimensional all liquid-phase method combined with MS for protein profile analysis was developed. Proteins are first fractionated by pI using IEF and then further separated according to hydrophobicity by reverse-phase HPLC (74). To improve its quantitative and high throughput capacity and its reproducibility, LC may be coupled to sample isotope labeling as in the ICAT technique (75). In this method, paired complex protein samples are isotope-labeled with tags, e.g. ^{12}C (light) and ^{13}C (heavy) that bind covalently to cysteine residues. Then samples are mixed, digested with trypsin, separated by HPLC, and identified by MS. The tags used are similar in structure and chemical properties but are different in mass. ICAT profiles the relative amounts of cysteine-containing peptides derived from tryptic digests of paired protein extracts (63). This method has been applied to compare NAFs from tumor-bearing and contralateral disease-free breasts of patients with unilateral early stage breast cancer (EBC), identifying and quantifying differences in various specific protein expressions (76). Another non-gel-based approach is multidimensional protein identification technology (MudPIT). Two steps of HPLC (cation exchange and reverse-phase) are coupled to MS/MS and database-searching algorithms allowing the rapid analysis of complex mixtures with direct identification of the generated peptide sequences. This technique has been recently associated to enzyme activity profiling in human tumor tissues, including breast tumors, and has generated functional signatures that correlate with previously described molecular subtypes (77).

One particular approach for the discovery of high quality protein signatures involves the identification of protein markers directly in tissue biopsy and straightforward analysis of a tissue section using MALDI-MS. This technique has been recently used for tumor samples, mainly lung and brain cancers, with good results in terms of diagnostic or prognostic classification (78, 79). Although only a few breast cancer samples have been analyzed so far (80), this technique represents a promising way to generate protein biomarkers from clinical samples.

SELDI-TOF MS is currently the most widely used and advertised non-gel-based method. The technique couples protein separation directly to presentation to the mass spectrometer. The first dimension of protein separation uses chromatographic substrates (the ProteinChip® array, Ciphergen Biosystems, Fremont, CA), which include anion exchange, cation exchange, normal phase, reverse-phase, and

IMAC. The various types of substrates have different affinity to different subsets of proteins, thereby increasing protein representation when combining various arrays. The combination of these arrays with up-front prefractionation using 96-well batch chromatography (e.g. anion exchange) allows the detection of up to 2,000 protein species from serum (81, 82). The resulting spectral masses are analyzed using univariate and multivariate statistical tools to yield a single marker or multi-marker pattern that can classify clinical samples. Discriminator protein peaks are then purified and submitted to the MS-based identification process. Knowledge of the chemistry that retains these potential biomarkers helps define the optimal purification process. Another approach may not require the identification of individual protein components but rather deals with discriminating patterns generated by sophisticated bioinformatic tools including artificial neural networks, genetic algorithms, and decision trees (83). The SELDI technique was developed to profile clinical biological fluids, notably serum/plasma, and got fame when numerous studies showed promising potential in identifying unique biomarkers or complex patterns with diagnostic value, allowing its use as a screening/early diagnosis tool in various cancers to be envisioned (84–88). One major criticism of the technique relies on the overall lack of sensitivity and capability to detect tumor-specific protein traces within a large amount of nonspecific protein species (89). This concern is particularly relevant in a screening or early diagnosis setting where tumor burden is expected to be minimal. However, although still controversial in its reproducibility and ability to detect actual specific tumor signatures, SELDI has several advantages, such as ease of use, high throughput, and relatively affordable cost, making it a very attractive technique for working with large sample cohorts in a clinical setting. The first results generated using this technique on clinical breast samples are discussed below.

CLINICAL APPLICATIONS IN BREAST CANCER RESEARCH

In this section, we focus on some studies that applied the two most popular proteomic technologies, TMAs and SELDI-TOF, to clinical issues in breast cancer.

Tissue Microarray Studies

Since the first publication in 1998 (23), breast cancer has been one of the most analyzed cancers using TMAs. Most studies focused on the clinical significance of a single or a few molecular markers individually on large series of samples. Recently tens of proteins were simultaneously tested in multivariate analysis to identify a multiprotein signature associated with a given phenotype. We present here some studies that addressed the diagnostic and prognostic issues.

Characterization of Previously Known Diagnostic Classes— Several studies compared tumor classes within different breast cancer classifications. The genetic classification distinguishes the sporadic forms (~95% of cases) and the he-

reditary forms (~5%); the latter are mainly due to mutations in *BRCA1* and *BRCA2* genes. Using DNA microarrays, Hedenfalk *et al.* (90) profiled breast cancers from *BRCA1* (seven samples) and *BRCA2* (eight samples) mutation carriers and identified a set of genes including *CCND1* that distinguished the two genotypes. Using a TMA that included 23 *BRCA1*- and 17 *BRCA2*-mutated tumors, overexpression of *CCND1* in *BRCA2* tumors was confirmed. The expression of 37 proteins was assessed in 20 *BRCA1*, 14 *BRCA2*, and 59 sporadic breast cancers on TMA (91). Proteins included hormone receptors, p53, ERBB2, cell cycle regulators, apoptosis markers, and basal markers. *BRCA1* and *BRCA2* tumors were different with respect to hormone receptors, cell cycle, and apoptosis and basal cell markers. The most striking difference was in the expression of cell cycle proteins with up-regulation of type D (D1 and D3) cyclins and their cyclin-dependent kinase (CDK4) regulators in *BRCA2* cancers with respect to *BRCA1* cancers. Most *BRCA1* tumors were also ER/ERBB2-negative and highly proliferative and had a basal phenotype.

The most frequent pathological type of breast cancer is ductal breast carcinoma (~80%). Medullary breast carcinoma (MBC) is a rare but enigmatic pathological type. Despite features of aggressiveness, it is associated with a favorable prognosis. Morphological diagnosis remains difficult, and very little is known about the molecular alterations involved in this disease. The expression of 18 proteins (markers of cellular origin, markers related to function and proliferation/differentiation of breast tissue, oncogenes, and tumor suppressors) was studied on 61 typical MBCs and 300 grade III non-MBCs (92). Typical MBCs were characterized by a high degree of basal/myoepithelial differentiation as shown with DNA microarrays (139). In multivariate analysis with logistic regression, typical MBC was optimally defined by high expression of P-cadherin and MIB1/Ki67, ERBB2 negativity, and p53 positivity.

Key proteins can be discovered by comparing the expression profiles of different stages of progression in mammary oncogenesis from normal to metastatic tissues. Expression of multifunctional 14-3-3 σ protein was studied in breast cancer (93). A series of 65 primary breast carcinomas was profiled by Western blotting, 2D electrophoresis, and IHC, providing evidence that, contrary to what was previously thought, 14-3-3 σ down-regulation at the protein level is not a frequent event in breast cancer. Results were validated by TMA on an independent set of 65 tumors. Another study used a “breast cancer progression TMA” including samples from 196 lymph node-negative tumors, 196 lymph node-positive tumors, and three different lymph node metastases from each node-positive tumor. The study demonstrated high concordance in ERBB2 status between primary tumors and their lymph node metastases (94).

Anatomoclinical forms of breast cancer include early forms (~85% of all cases), locally advanced forms (~10%), and metastatic forms (~5%). A rare (~3%) but often lethal form is

inflammatory breast cancer (IBC). Diagnosis, based on clinical and/or pathological criteria, may be difficult, and little is known about the underlying molecular alterations. Two TMA studies compared the protein profile of IBCs and non-IBCs (NIBCs). The first study profiled 34 IBCs and 41 NIBCs and confirmed the overexpression of E-cadherin and RhoC GTPase in IBC (95). We analyzed the expression of eight proteins (ER, progesterone receptor, E-cadherin, EGFR, ERBB2, MUC1, MIB1, and p53) in 80 IBCs and 552 NIBCs (96). All proteins were differentially expressed between IBCs and NIBCs in univariate analysis; five protein expressions were significant in multivariate analysis: high expression of E-cadherin, ER negativity, MIB1 positivity, MUC1 cytoplasmic staining, and ERBB2 positivity. The probability that a tumor with the complete protein signature was an IBC was 91% but less than 50% when three or fewer parameters were present.

Molecular Identification of New Breast Cancer Subclasses—Gene expression studies and clustering methods have defined biologically and clinically relevant subclasses of breast cancer. In pioneering studies, global hierarchical clustering partitioned samples in two major groups, ER-positive and ER-negative, indicating that many genes display expression positively or negatively associated to ER status (5, 97–99). Perou *et al.* (100) refined this classification by making reference to the epithelial cell types. By confronting expression profiles of tumors and cell lines, they identified gene clusters associated to cell types (such as luminal and basal) and others specific to genomic alterations or signaling pathways (such as ERBB2 cluster). These connections were corroborated using IHC with antibodies against luminal and basal cell cytokeratins. Subsequent analyses refined this model with the subdivision of ER-positive tumors, resulting in the definition of at least five molecular subtypes (luminal A, luminal B, basal, ERBB2-overexpressing, and normal breastlike), which correlated with relevant histoclinical data (3, 10–12).

Six proteomic studies attempted to validate this classification on paraffin-embedded tumors at the protein level. Clustering was applied to IHC data of 7–31 proteins in large series of samples, ranging from 97 to 1,944 patients. Most of the markers tested had a well established or suspected role in breast cancer or represented proteins for which the transcript was a discriminator in previous gene profiling studies. ER, progesterone receptor, ERBB2, and p53 were present in all studies; CK5/6 was present in five of six; CK8/18, cyclin E, and Ki67 were present in four; and BCL2, cyclin D1, E-cadherin, and other ERBBs were present in three studies. Hierarchical clustering revealed concordant protein clusters and tumor clusters in agreement with RNA expression studies with clear distinction between ER-negative and ER-positive tumors and existence of subgroups in each category. The biological significance of these clusters was supported by their similarity with those derived from DNA microarrays studies, the confirmation of known correlations between proteins, and the presence of luminal and basal epithelial markers. Korsch-

ing *et al.* (101) identified luminal and basal tumor clusters without any histoclinical information. Others (102, 103) confirmed the known correlation of molecular subtypes with pathological characteristics. These first studies also indicated the potential of clustering applied to qualitative or semiquantitative data for tumor grouping. Three additional studies reported correlations with survival. Using IHC on TMAs, we monitored the expression of 26 proteins in tumor samples from 552 consecutive EBC patients (104). With a median follow-up of ~5 years, the 5-year MFS was 80%. The 26 proteins included hormone receptors, subclass markers, oncogenes and proliferation proteins, tumor suppressors, adhesion molecules, and proteins coded by amplified genes. Clustering delineated four protein clusters related to metabolic pathway (“ER cluster”) and functions (“proliferation,” “mitosis,” and “differentiation” clusters) and sorted tumors in three clusters (A1, A2, and B, including, respectively, 471, 62, and 81 samples). The profiles suggested that A1 tumors may be approximated to luminal-like tumors and that B tumors may be approximated to basal-like tumors. Tumor clusters correlated with grade, ER and ERBB2 status, and survival with 5-year MFS of 86% in A1, 62% in A2, and 64% in B ($p < 0.0001$). Although ER expression was a key factor in our classification, ER-positive samples and ER-negative samples displayed heterogeneous expression profiles with the identification of at least two subgroups with different survival in each category as previously reported in gene expression studies. This suggested that the grouping of tumors based on the expression of multiple proteins, including ER, was more powerful than ER alone to tackle the heterogeneity of disease. In agreement, the clustering-based grouping was the most significant prognostic factor in multivariate analysis. Similar results were reported for 438 breast cancers profiled with 31 markers and 1,944 tumors profiled with 25 proteins (105, 106). Clearly these studies showed that clustering of IHC data by using a few tens of markers can identify biologically and clinically relevant subtypes of disease more reliably than an individual prognostic marker.

Proteomic studies thus reinforced the robustness of the molecular taxonomy that emerged from DNA microarray analyses. Currently luminal and ERBB2-positive tumors have clinical diagnostic assays (IHC for ER and IHC and fluorescence *in situ* hybridization for ERBB2) and molecular therapeutic targets (hormone therapy and trastuzumab, respectively). Conversely the basal group is poorly characterized with neither validated diagnostic assay nor therapeutic target. Two studies used TMAs and IHC to better characterize basal samples and to identify a specific protein pattern (107, 108). Nielsen *et al.* (107) collected 21 tumors defined as basal by gene profiling and measured the expression of five proteins encoded by genes included in or negatively correlated with the basal gene cluster (ER, ERBB2, EGFR, CK5/6, and KIT). The pattern defined by ER-negative, ERBB2-negative, CK5/6-positive, and/or EGFR-positive was 76% sensitive and

100% specific for identifying the 21 basal tumors. Tested on a TMA containing 663 breast cancers with long follow-up, such a pattern delineated a group of tumors (15% of cases) associated with a much worse survival than the two other groups (ER-positive/ERBB2-negative and ERBB2-positive).

Prognostic Studies—Most studies on breast cancer prognosis using TMAs analyzed a single protein, whereas only two studies analyzed tens of proteins to identify a multiprotein signature. Examples of single tested protein with previously known or suspected relevance in breast cancer are multiple. In a series of 1,576 tumors (109), elevated expression of COX2 was associated with unfavorable outcome. Other proteins with negative prognostic impact included the catalytic subunit of telomerase (TERT) tested in 611 tumors (110), the epithelial cellular adhesion molecule (Ep-CAM/TACSTD1) in 1,715 tumors (111), or the early placenta insulin-like growth factor (EPIL) in 603 tumors (112). Expression of KIT, a target of imatinib therapy, was not associated with survival in a series of 1,654 tumors (113). Conversely proteins associated with favorable clinical outcome included fragile histidine triad (114), tyrosine-phosphorylated STAT5 (115), or BCL2 (116). Thirteen proteins were first tested in 930 tumors (116). BCL2 was the only independent prognostic factor in multivariate analysis combining IHC and the Nottingham Prognostic Index. BCL2 impact was then confirmed in an independent series of 1,961 tumors.

TMAs have also been used to validate and extend large series of sample results from gene profiling studies. One of the most prominent genes within the ER-related luminal gene cluster is the transcription factor GATA3. The favorable prognostic impact of GATA3 expression was confirmed at the protein level in 552 (104) and 139 tumors (117). Genes included in the basal cluster were also tested, and their negative prognostic role was confirmed in several hundreds of tumors: crystallin alpha B (118), cytokeratins 5 and 17 (119), and annexin A8 (120). Using DNA microarrays, we identified a gene expression signature associated with ERBB2 status (121). We then used TMAs (~250 tumors) to validate at the protein level the positive correlation of GATA4 and Ki67 with ERBB2. Similarly Rouzier *et al.* (9) confirmed on 122 independent tumors the correlation between low expression of the microtubule-associated protein TAU and sensitivity to neoadjuvant paclitaxel identified by gene profiling. Identification of a multiprotein signature of prognosis by supervised analysis, as reported at the RNA level with DNA microarrays, is more challenging with IHC data in terms of statistics. It has been reported so far in two studies. We developed a supervised method to search for the best combination within 26 tested proteins that would improve the prognostic classification of 552 tumors (104). We identified a 21-protein set (Fig. 1A) that optimally classified patients in two classes (“good prognosis” and “poor prognosis”) with different 5-year MFS. Initially identified in a learning set of 368 patients, this prognostic signature was validated in an independent set of 184 patients,

showing its robustness. The classification based on the 21-protein predictor was associated with a difference in 5-year MFS (90 versus 61%; Fig. 1B). This difference remained significant when molecular grouping was done according to lymph node invasion, ER status, or type of adjuvant systemic therapy. Our predictor performed well in patients irrespective of ER status, suggesting it provides more accurate clinical information than ER status alone, possibly reflecting functional differences in the ER pathway or interacting pathways. In multivariate analysis, which included other prognostic factors and each protein separately, the 21-protein set was the strongest independent predictor of clinical outcome. This study was the first to apply both unsupervised and supervised analyses to qualitative IHC data. More recently, supervised analysis was applied to nine proteins and 324 stage I to III breast cancers treated with adjuvant tamoxifen (122). Univariate analysis showed the importance of conditional interpretation of certain markers on others due to their interdependency. In combination with the application of sophisticated informatics, it enabled the development of a predictive model that performed better than the current standards.

SELDI-TOF Studies

In breast cancer, SELDI-TOF MS was used to investigate serum/plasma (123–126) but also NAF (127–130) and tumor tissues (131) as a potential source for diagnostic biomarkers and recently as a potential tool to predict outcome (131, 132) and/or to monitor cancer treatment (133, 134).

SELDI and Diagnosis—Serum/plasma is the most easily available sample type for clinical proteomic studies. It may be a reliable surrogate tissue for a physiologic or pathologic process such as cancer. Some studies have addressed the potential of SELDI to provide serum biomarkers that differentiate breast cancer from benign disease and/or healthy subjects (123, 124, 126). Enrolling between 133 and 169 patients, these studies have identified diagnostic protein profiles with sensitivities and specificities of 76–93 and 90–93%, respectively. Protein peaks that distinguished healthy women from those with cancer were found at *m/z* 4,300 and 8,900 in two studies. However, no protein identification was provided. Another serum-based study with diagnostic purpose (125) analyzed 30 serum samples from BRCA1 mutation patients who either developed cancer ($n = 15$) or remained cancer-free ($n = 15$). It revealed 23 specific markers in BRCA1 cancer specimens that classified the two groups with 87% sensitivity and specificity. This study also included a comparison of serum profiles of BRCA1 mutation cancer developers and patients with sporadic breast cancers. Various proteins were up-regulated in the BRCA1 specimens and classified the two groups with 94% sensitivity and 100% specificity. Again no protein identity was reported, and the limited sample size precluded independent validation.

Other SELDI-based diagnostic studies used materials more

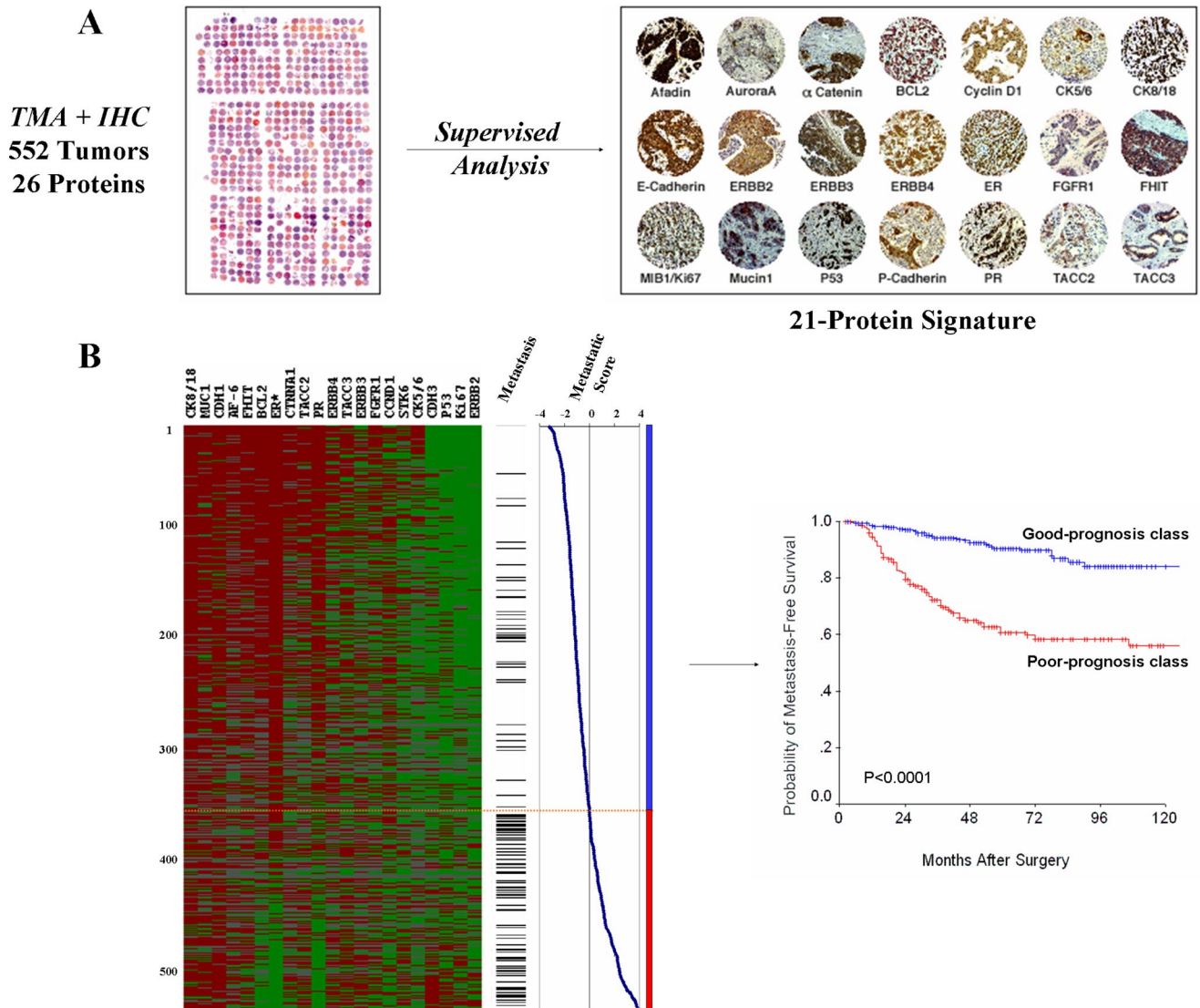


FIG. 1. Identification of a prognostic multiprotein signature in breast cancer samples using IHC and TMA. *A, left*, representative hematoxylin-eosin and safran staining of a paraffin block section ($25 \times 30 \text{ mm}^2$) from a TMA containing 552 early breast cancer cases with 0.6-mm tumor cores. Using IHC on TMAs, we monitored the expression of 26 proteins in tumor samples from 552 consecutive early breast cancer patients. *Right*, immunohistochemical staining of a tumor core for the 21 proteins included in the prognostic signature identified by supervised analysis (magnification, $\times 200$). *B, left*, supervised classification of 552 samples using the 21-protein expression signature. Each row of the data matrix (*left panel*) represents a sample, and each column represents a protein. Immunostaining results are depicted according to a color scale: *red* for positive and *green* for negative staining. The 21 proteins, listed above the matrix (*ER**, means of two independent ER analyses), are ordered from *left to right* according to decreasing Δp (Δp is the difference between the probability of positive staining and the probability of negative staining in non-metastatic samples). Tumor samples are numbered from 1 to 552 and are ordered from *top to bottom* according to their increasing “metastatic score” (*right panel*). The *orange dashed line* indicates the threshold 0 that separates the two classes of samples, poor prognosis (*under the line, red class*) and good prognosis (*above the line, blue class*). The *middle panel* indicates the occurrence (*black square*) or not (*white square*) of metastatic relapse for each patient. *Right*, Kaplan-Meier analysis of the metastasis-free survival according to the 21-protein expression signature-based classification. The p value is calculated using the log-rank test (adapted from Ref. 104). *PR*, progesterone receptor; *FHIT*, fragile histidine triad.

directly related to the tumor such as NAF or ductal lavage (128–130, 135–138). These studies, most of them with limited sample size, compared NAF or ductal lavage from healthy subjects (including benign breast disease for some) with unrelated patients with cancer or examined matched pair samples from women with unilateral primary invasive breast can-

cer using the contralateral (unaffected) breast as the “normal” sample. Considerable interpatient heterogeneity was reported, whereas similarity between matched NAF from cancerous and unaffected breast was high (128, 136). However, protein peaks with differential expression between cancerous versus healthy NAF were observed, some of them correlating

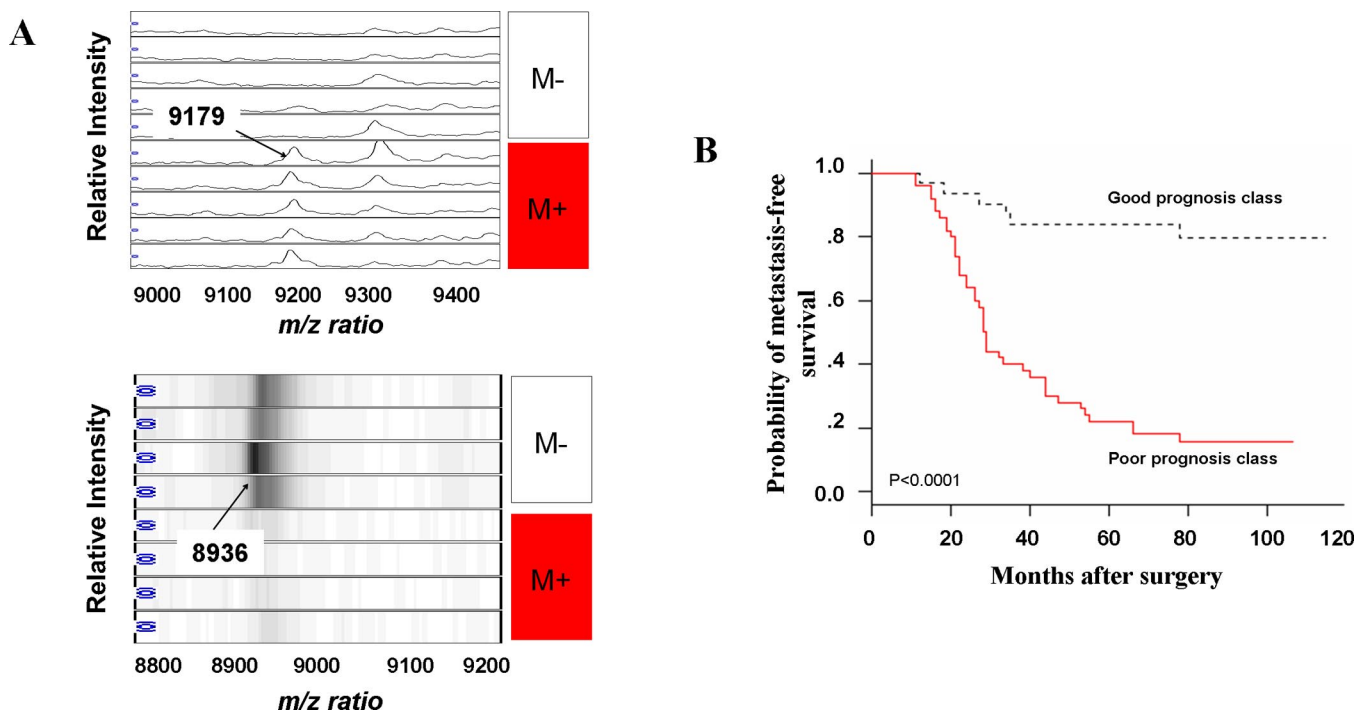


FIG. 2. Identification of a prognostic multiprotein signature in serum from early breast cancer patients treated with adjuvant chemotherapy using SELDI-TOF profiling. *A*, top, a protein with m/z of 9,179 (spectra view) is up-regulated in patients with metastatic outcome ($M+$) compared with metastasis-free surviving patients ($M-$). Bottom, a protein with m/z of 8,936 is down-regulated in patients with metastatic outcome ($M+$) compared with metastasis-free surviving patients ($M-$) (gel view). *B*, Kaplan-Meier analysis of the metastasis-free survival according to the serum multiprotein-based classification.

with axillary lymph node metastases and extent of the disease (135). One of the largest studies compared NAF from 27 cancer patients with 87 control women and differentiated invasive cancer not only from healthy subjects but also from ductal carcinoma *in situ* and benign breast disease (138).

SELDI and Prognosis—Whereas a large number of SELDI-based diagnostic studies were rapidly reported, little is known about the potential of this technique for prognostic evaluation. Using SELDI-TOF MS, we have reported the first evidence that serum protein profiling can provide prognosis information in EBC. We analyzed postoperative early serum from a retrospective series of 81 EBC patients receiving adjuvant anthracycline-based chemotherapy (132). Sera collected after surgery and before any specific adjuvant treatment were fractionated by combining chromatographic anion exchange beads and retention on chromatographic protein chips and analyzed by TOF-based mass spectrometry. Proteins differentially expressed (Fig. 2A) according to the metastatic outcome were selected and subjected to statistical analysis combining supervised partial least squares projection and a logistic regression model. A 40-protein index was generated that correctly predicted the clinical outcome in 83% of patients and was validated using leave-one-out cross-validation. This protein index identified two classes of patients with a difference in 5-year MFS and overall survival (Fig. 2B). It was the strongest independent prognostic factor in multivariate

analysis, which included the clinicopathological factors with prognostic impact in univariate analysis, such as lymph node invasion, pathological tumor size, grade, and hormonal receptivity. Some of the components of the protein index were identified: haptoglobin α -1 and transferrin levels were up-regulated, whereas C3a complement fraction was down-regulated in serum from patients with high probability of metastatic relapse. Other identified proteins included apolipoproteins A-I and C-I. All these proteins are non-tumor-specific and are most certainly derived from host response. However, some of them are involved in biological processes such as angiogenesis and immune response, which may influence metastasis outcome. Although still preliminary and requiring validation on an independent dataset, these results provide incentive to further explore SELDI-based serum proteomics as a prognostic and/or predictive tool.

Another study has recently investigated the ability of SELDI technology to address prognostic issues (131). This study profiled tumor cytosolic extracts from 60 patients with node-negative EBC (most of them without adjuvant chemotherapy), 30 of them displaying a metastatic relapse and 30 of them being disease-free after the follow-up period. Differentially expressed proteins according to outcome were examined, and two protein peaks, proved by univariate Cox regression to have a good cumulative prognostic-informative power, improved the conventional Saint-Gallen prognostic classifica-

tion of node-negative breast cancer (140). These two peaks were purified and identified by LDI-Qq-TOF MS/MS as ubiquitin and ferritin light chain; their up-regulation was confirmed in tumor tissues by Western blotting analyses and IHC.

SELDI and Treatment Monitoring—Another promising clinical application of SELDI profiling is the identification of biomarkers that could predict clinical benefit and/or toxicities of treatment. Access to serum allows easy serial sampling in clinical settings. Early alteration of serum protein patterns may be detected that could be associated to toxic side effects and/or therapeutic response. Two studies addressed these issues in breast cancer patients using SELDI-based approaches. Plasma samples from newly diagnosed stage I–III breast cancer patients receiving neoadjuvant or adjuvant chemotherapy were profiled (134). By comparing post-treatment day 3 to pretreatment samples, in both settings, a similar increase in a single peak at 2.7 kDa was found. This alteration seemed to be induced by paclitaxel and was much less pronounced in anthracycline-treated patients. Serum pharmacoproteomic profiling was also done in five docetaxel-treated breast cancer patients from 4 to 48 h following drug infusion (133). A similar expression pattern of two proteins with m/z of 7,700 and 9,200 (subsequently identified as high molecular weight kininogen and apolipoprotein A-II, respectively) was found in four patients. This pattern was not observed in the fifth patient who experienced severe and acute adverse effects.

Another field of interest for SELDI-based profiling of human breast cancer samples may be the prediction of therapeutic response to anticancer agents when applied to more advanced tumors. Compared with early diagnosis-oriented studies, this approach (which can be named “theragnostics”) may be more easily successful. Large tumor burden is frequently present in this setting, increasing the theoretical concentration of putative tumor markers that could regulate drug activity and therefore the probability of their detection in body fluids. In our institution, we examine SELDI-generated serum and tumor profiles from metastatic breast cancer patients receiving specific anticancer treatment (trastuzumab-based in ERBB2-positive patients and docetaxel-based in ERBB2-negative patients). Discriminator proteins between responder and non-responder patients might be identified.

PERSPECTIVES

Altogether these preliminary studies have revealed the potential of proteomics in clinical issues of breast cancer both to validate results from high throughput molecular studies and to define new potential diagnostic, prognostic, and predictive protein signatures. Results (obtained on relatively small number of samples for non-TMA-based studies) must now be validated, refined, and extended. Several obstacles have to be addressed to allow proteomics to reach its optimal yield and to translate into tangible benefits for patients. The accessibility, relatively limited to some academic facilities or com-

panies during the past years, is increasing rapidly, notably for TMAs, and commercial TMAs are now available. The availability, complexity, and cost of other proteomic tools will likely improve with the next generation of tools. Because of the complexity and high throughput of experiments, exquisite quality controls are required, including controlled experimental conditions, signal-to-noise ratio, dynamic and linear range, sensitivity, and reproducibility. With regard to SELDI, identification of the discriminator protein(s) may be problematic. A major issue is the handling of data for multiprotein experiments. Collaborations are being established between researchers across disciplines (physicians, biologists, mathematicians, statisticians, and computer scientists) to develop adequate tools required for producing, storing, analyzing, visualizing, and interpreting data. Previous works on gene expression profiling data constitute precious and useful bases for progresses, notably for statistical analyses (141). However, these methods do not easily translate TMA data, and the development of more specific methods is needed. Efforts are required to make results publicly available as done with gene profiling data. This is underway for TMA data: for example the tissue microarray data exchange specification developed by the Association of Pathology Informatics and the NCI, National Institutes of Health, allows researchers to submit their data to journals and to public data repositories and to share or merge data from different laboratories (142). Such effort will facilitate the comparison of different experimental procedures and results and the development of new analytic tools.

Other issues are related to samples. A major issue is the all too pressing urge to have a high quality sample bank linked to a searchable database containing all histoclinical parameters of tumors as well as patient consent and confidentiality. The careful and uniform collection of samples and data should become a component of all future clinical trials. Except for IHC on TMA that works well on FFPE samples, good preservation of frozen samples is crucial because the quality of data is sensitive to protein degradation. The amount of material required for experiments is an important issue because of the small size of clinical specimens and the absence of techniques of amplification. Improvements of microtechniques and nanotechnologies may soon provide ways to circumvent this obstacle. The development of kits that can extract RNA and proteins from frozen samples simultaneously is also urgently needed. The heterogeneity of cancer tissue samples is another potential source of variability of proteomic data. Breast tumors contain several malignant clones with different metastatic capacity. Furthermore tumor tissue samples contain several cell types in different proportions. This issue may be addressed by confronting expression profiles of macrodissected samples with those of lines representing the different cell types present (“virtual microdissection”). A different and far more labor-intensive solution lies in the use of laser microdissection, which allows the harvesting of pure cell subpopulations from frozen or fixed tissue (143). Actually perhaps

the most attractive material that could be analyzed by proteomics in the near future would be cancer stem cells. Breast cancer stem cells have been identified (43). They are thought to fuel the tumor mass and should be the cells to target by specific therapy.

Finally once a potential marker has been identified by proteomic analysis and before any clinical application, it needs to be validated in a large series of samples. In this context, IHC on TMA is the most interesting approach provided that the appropriate antibody is available or can be made. It is too early to define the format of proteomic tool that will be applied in clinical practice. One example includes protein arrays that could detect, in the tumor and among multiple kinases, the activation of kinases to be targeted by specific therapies.

CONCLUSION

Proteomics of breast cancer is still in its infancy, but there is no doubt that it has a brilliant future for several reasons. First, it allows identification of targets for specific therapy. Proteins are excellent targets when mutated, truncated, or expressed as specific isoforms or in large amount (e.g. after gene amplification). Cancer isoforms are currently found after genomic studies rather than protein studies (e.g. DNA sequencing and molecular cytogenetics), but this could change with the advent of new techniques to study tumor samples and fluids. Second, proteins are good markers. Tumors and patient fluids may express specific protein isoforms that may serve as beacons for early detection and/or follow-up. Third, proteomics may be easily coupled with functional tests. Antibody arrays or cell arrays are good examples of this. Fourth, in contrast to other levels of information, proteomics may be applied to components of the cell, such as nucleus, membrane, and organelles, which may deserve specific attention or be more convenient to study in some cases. Even if, as mentioned above, a number of issues remain to be addressed before the proteomics can enter clinical practice, it is anticipated that in the future such an approach will affect the current generation of physicians in charge of breast cancer patients, hopefully transforming disease management into a more structured and logical science and a successful medicine. The characterization of discriminator proteins will certainly provide new markers useful for screening, diagnosis, prognosis, and monitoring and will help in developing new molecularly targeted anticancer drugs.

* This work was supported by Institut Paoli-Calmettes, INSERM, Université de la Méditerranée, and grants from the Ministries of Health and Research (Cancéropôle ACI2004 and ACI2005) and Ligue Nationale contre le Cancer (label). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

||To whom correspondence should be addressed: Dépt. d'Oncologie Moléculaire, Inst. Paoli-Calmettes, 232 Bd. Sainte-Marguerite, 13009 Marseille, France. Tel.: 33-4-91-22-35-37; Fax: 33-4-91-22-36-70; E-mail: bertucci@marseille.fnclcc.fr.

REFERENCES

- Bertucci, F., Houlgatte, R., Nguyen, C., Viens, P., Jordan, B. R., and Birnbaum, D. (2001) Gene expression profiling of cancer by use of DNA arrays: how far from the clinic? *Lancet Oncol.* **2**, 674–682
- Ayers, M., Symmans, W. F., Stec, J., Damokosh, A. I., Clark, E., Hess, K., Lecoche, M., Metivier, J., Booser, D., Ibrahim, N., Valero, V., Royce, M., Arun, B., Whitman, G., Ross, J., Sneige, N., Hortobagyi, G. N., and Puzstai, L. (2004) Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J. Clin. Oncol.* **22**, 2284–2293
- Bertucci, F., Finetti, P., Rougemont, J., Charafe-Jauffret, E., Cervera, N., Tarpin, C., Nguyen, C., Xerri, L., Houlgatte, R., Jacquemier, J., Viens, P., and Birnbaum, D. (2005) Gene expression profiling identifies molecular subtypes of inflammatory breast cancer. *Cancer Res.* **65**, 2170–2178
- Bertucci, F., Finetti, P., Rougemont, J., Charafe-Jauffret, E., Nasser, V., Lloriod, B., Camerolo, J., Tagett, R., Tarpin, C., Houvenaeghel, G., Nguyen, C., Maraninchi, D., Jacquemier, J., Houlgatte, R., Birnbaum, D., and Viens, P. (2004) Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. *Cancer Res.* **64**, 8558–8565
- Bertucci, F., Houlgatte, R., Benziane, A., Granjeaud, S., Adelaide, J., Tagett, R., Lloriod, B., Jacquemier, J., Viens, P., Jordan, B., Birnbaum, D., and Nguyen, C. (2000) Gene expression profiling of primary breast carcinomas using arrays of candidate genes. *Hum. Mol. Genet.* **9**, 2981–2991
- Bertucci, F., Nasser, V., Granjeaud, S., Eisinger, F., Adelaide, J., Tagett, R., Lloriod, B., Giaconia, A., Benziane, A., Devillard, E., Jacquemier, J., Viens, P., Nguyen, C., Birnbaum, D., and Houlgatte, R. (2002) Gene expression profiles of poor-prognosis primary breast cancer correlate with survival. *Hum. Mol. Genet.* **11**, 863–872
- Chang, J. C., Wooten, E. C., Tsimelzon, A., Hilsenbeck, S. G., Gutierrez, M. C., Elledge, R., Mohsin, S., Osborne, C. K., Chamness, G. C., Allred, D. C., and O'Connell, P. (2003) Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* **362**, 362–369
- Ma, X. J., Wang, Z., Ryan, P. D., Isakoff, S. J., Barmettler, A., Fuller, A., Muir, B., Mohapatra, G., Salunga, R., Tuggle, J. T., Tran, Y., Tran, D., Tassin, A., Amon, P., Wang, W., Enright, E., Stecker, K., Estepa-Sabal, E., Smith, B., Younger, J., Balis, U., Michaelson, J., Bhan, A., Habin, K., Baer, T. M., Brugge, J., Haber, D. A., Erlander, M. G., and Sgroi, D. C. (2004) A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* **5**, 607–616
- Rouzier, R., Perou, C. M., Symmans, W. F., Ibrahim, N., Cristofanilli, M., Anderson, K., Hess, K. R., Stec, J., Ayers, M., Wagner, P., Morandi, P., Fan, C., Rabiul, I., Ross, J. S., Hortobagyi, G. N., and Puzstai, L. (2005) Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin. Cancer Res.* **11**, 5678–5685
- Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Eystein Lonning, P., and Borresen-Dale, A. L. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874
- Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C. M., Lonning, P. E., Brown, P. O., Borresen-Dale, A. L., and Botstein, D. (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8418–8423
- Sotiriou, C., Neo, S. Y., McShane, L. M., Korn, E. L., Long, P. M., Jazaeri, A., Martiat, P., Fox, S. B., Harris, A. L., and Liu, E. T. (2003) Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10393–10398
- van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E. T., Friend, S. H., and Bernards, R. (2002) A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009

14. Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M. E., Yu, J., Jatkoe, T., Berns, E. M., Atkins, D., and Foekens, J. A. (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **365**, 671–679
15. Anderson, L., and Seilhamer, J. (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18**, 533–537
16. Chen, G., Gharib, T. G., Wang, H., Huang, C. C., Kuick, R., Thomas, D. G., Shedden, K. A., Misek, D. E., Taylor, J. M., Giordano, T. J., Kardia, S. L., Iannettoni, M. D., Yee, J., Hogg, P. J., Orringer, M. B., Hanash, S. M., and Beer, D. G. (2003) Protein profiles associated with survival in lung adenocarcinoma. *Proc. Natl. Acad. Sci. U. S. A* **100**, 13537–13542
17. Ginestier, C., Charafe-Jauffret, E., Bertucci, F., Eisinger, F., Geneix, J., Bechlian, D., Conte, N., Adelaide, J., Toiron, Y., Nguyen, C., Viens, P., Mozziconacci, M. J., Houlgatte, R., Birnbaum, D., and Jacquemier, J. (2002) Distinct and complementary information provided by use of tissue and DNA microarrays in the study of breast tumor markers. *Am. J. Pathol.* **161**, 1223–1233
18. Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730
19. Nishizuka, S., Charboneau, L., Young, L., Major, S., Reinhold, W. C., Waltham, M., Kourou-Mehr, H., Bussey, K. J., Lee, J. K., Espina, V., Munson, P. J., Petricoin, E., III, Liotta, L. A., and Weinstein, J. N. (2003) Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc. Natl. Acad. Sci. U. S. A* **100**, 14229–14234
20. Tyers, M., and Mann, M. (2003) From genomics to proteomics. *Nature* **422**, 193–197
21. Battifora, H. (1986) The multitumor (sausage) tissue block: novel method for immunohistochemical antibody testing. *Lab. Invest.* **55**, 244–248
22. Wan, W. H., Fortuna, M. B., and Furmanski, P. (1987) A rapid and efficient method for testing immunohistochemical reactivity of monoclonal antibodies against multiple tissue samples simultaneously. *J. Immunol. Methods* **103**, 121–129
23. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat. Med.* **4**, 844–847
24. Schraml, P., Kononen, J., Bubendorf, L., Moch, H., Bissig, H., Nocito, A., Mihatsch, M. J., Kallioniemi, O. P., and Sauter, G. (1999) Tissue microarrays for gene amplification surveys in many different tumor types. *Clin. Cancer Res.* **5**, 1966–1975
25. Lugli, A., Spichtin, H., Maurer, R., Mirlacher, M., Kiefer, J., Huusko, P., Azorsa, D., Terracciano, L., Sauter, G., Kallioniemi, O. P., Mousseis, S., and Tornillo, L. (2005) EphB2 expression across 138 human tumor types in a tissue microarray: high levels of expression in gastrointestinal cancers. *Clin. Cancer Res.* **11**, 6450–6458
26. Went, P. T., Dirnhofer, S., Bundi, M., Mirlacher, M., Schraml, P., Mangialaio, S., Dimitrijevic, S., Kononen, J., Lugli, A., Simon, R., and Sauter, G. (2004) Prevalence of KIT expression in human tumors. *J. Clin. Oncol.* **22**, 4514–4522
27. Oode, K., Furuya, T., Harada, K., Kawauchi, S., Yamamoto, K., Hirano, T., and Sasaki, K. (2000) The development of a cell array and its combination with laser-scanning cytometry allows a high-throughput analysis of nuclear DNA content. *Am. J. Pathol.* **157**, 723–728
28. Richter, J., Wagner, U., Kononen, J., Fijan, A., Bruderer, J., Schmid, U., Ackermann, D., Maurer, R., Alund, G., Knogel, H., Rist, M., Wilber, K., Anabitarte, M., Hering, F., Hardmeier, T., Schonenberger, A., Flury, R., Jager, P., Fehr, J. L., Schraml, P., Moch, H., Mihatsch, M. J., Gasser, T., Kallioniemi, O. P., and Sauter, G. (2000) High-throughput tissue microarray analysis of cyclin E gene amplification and overexpression in urinary bladder cancer. *Am. J. Pathol.* **157**, 787–794
29. Kocher, T., Zheng, M., Bolli, M., Simon, R., Forster, T., Schultz-Thater, E., Rimmel, E., Noppen, C., Schmid, U., Ackermann, D., Mihatsch, M. J., Gasser, T., Heberer, M., Sauter, G., and Spagnoli, G. C. (2002) Prognostic relevance of Mage-A4 tumor antigen expression in transitional cell carcinoma of the urinary bladder: a tissue microarray study. *Int. J. Cancer* **100**, 702–705
30. Bubendorf, L., Kononen, J., Koivisto, P., Schraml, P., Moch, H., Gasser, T. C., Willi, N., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1999) Survey of gene amplifications during prostate cancer progression by high-throughput fluorescence in situ hybridization on tissue microarrays. *Cancer Res.* **59**, 803–806
31. Bubendorf, L., Kolmer, M., Kononen, J., Koivisto, P., Mousseis, S., Chen, Y., Mahlamaki, E., Schraml, P., Moch, H., Willi, N., Elkahloun, A. G., Pretlow, T. G., Gasser, T. C., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1999) Hormone therapy failure in human prostate cancer: analysis by complementary DNA and tissue microarrays. *J. Natl. Cancer Inst.* **91**, 1758–1764
32. Bertucci, F., Salas, S., Eysteris, S., Nasser, V., Finetti, P., Ginestier, C., Charafe-Jauffret, E., Lloriod, B., Bachelart, L., Montfort, J., Victorero, G., Viret, F., Ollendorff, V., Fert, V., Giovaninni, M., Delpero, J. R., Nguyen, C., Viens, P., Monges, G., Birnbaum, D., and Houlgatte, R. (2004) Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. *Oncogene* **23**, 1377–1391
33. Chung, G. G., Provost, E., Kielhorn, E. P., Charette, L. A., Smith, B. L., and Rimm, D. L. (2001) Tissue microarray analysis of beta-catenin in colorectal cancer shows nuclear phospho- β -catenin is associated with a better prognosis. *Clin. Cancer Res.* **7**, 4013–4020
34. Moch, H., Schraml, P., Bubendorf, L., Mirlacher, M., Kononen, J., Gasser, T., Mihatsch, M. J., Kallioniemi, O. P., and Sauter, G. (1999) High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. *Am. J. Pathol.* **154**, 981–986
35. Hoos, A., and Cordon-Cardo, C. (2001) Tissue microarray profiling of cancer specimens and cell lines: opportunities and limitations. *Lab. Invest.* **81**, 1331–1338
36. Brazma, A., and Vilo, J. (2000) Gene expression data analysis. *FEBS Lett.* **480**, 17–24
37. Liu, C. L., Prapong, W., Natkunam, Y., Alizadeh, A., Montgomery, K., Gilks, C. B., and van de Rijn, M. (2002) Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. *Am. J. Pathol.* **161**, 1557–1565
38. Alkushi, A., Irving, J., Hsu, F., Dupuis, B., Liu, C. L., Rijn, M., and Gilks, C. B. (2003) Immunoprofile of cervical and endometrial adenocarcinomas using a tissue microarray. *Virchows Arch.* **442**, 271–277
39. Rhodes, D. R., Sanda, M. G., Otte, A. P., Chinnaiyan, A. M., and Rubin, M. A. (2003) Multiplex biomarker approach for determining risk of prostate-specific antigen-defined recurrence of prostate cancer. *J. Natl. Cancer Inst.* **95**, 661–668
40. Alonso, S. R., Ortiz, P., Pollan, M., Perez-Gomez, B., Sanchez, L., Acuna, M. J., Pajares, R., Martinez-Tello, F. J., Hortelano, C. M., Piris, M. A., and Rodriguez-Peralto, J. L. (2004) Progression in cutaneous malignant melanoma is associated with distinct expression profiles: a tissue microarray-based study. *Am. J. Pathol.* **164**, 193–203
41. Nielsen, T. O., Hsu, F. D., O'Connell, J. X., Gilks, C. B., Sorensen, P. H., Linn, S., West, R. B., Liu, C. L., Botstein, D., Brown, P. O., and Van De Rijn, M. (2003) Tissue microarray validation of epidermal growth factor receptor and SALL2 in synovial sarcoma with comparison to tumors of similar histology. *Am. J. Pathol.* **163**, 1449–1456
42. Torhorst, J., Bucher, C., Kononen, J., Haas, P., Zuber, M., Kochli, O. R., Mross, F., Dieterich, H., Moch, H., Mihatsch, M., Kallioniemi, O. P., and Sauter, G. (2001) Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am. J. Pathol.* **159**, 2249–2256
43. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., and Clarke, M. F. (2003) Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3983–3988
44. Camp, R. L., Chung, G. G., and Rimm, D. L. (2002) Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat. Med.* **8**, 1323–1327
45. Camp, R. L., Dolled-Filhart, M., King, B. L., and Rimm, D. L. (2003) Quantitative analysis of breast cancer tissue microarrays shows that both high and normal levels of HER2 expression are associated with poor outcome. *Cancer Res.* **63**, 1445–1448
46. Harigopal, M., Berger, A. J., Camp, R. L., Rimm, D. L., and Kluger, H. M. (2005) Automated quantitative analysis of E-cadherin expression in lymph node metastases is predictive of survival in invasive ductal breast cancer. *Clin. Cancer Res.* **11**, 4083–4089
47. Rubin, M. A., Zerkowski, M. P., Camp, R. L., Kuefer, R., Hofer, M. D., Chinnaiyan, A. M., and Rimm, D. L. (2004) Quantitative determination of expression of the prostate cancer protein alpha-methylacyl-CoA race-

- mase using automated quantitative analysis (AQUA): a novel paradigm for automated and continuous biomarker measurements. *Am. J. Pathol.* **164**, 831–840
48. Giltmane, J. M., and Rimm, D. L. (2004) Technology insight: identification of biomarkers with tissue microarray technology. *Nat. Clin. Pract. Oncol.* **1**, 104–111
 49. Haab, B. B. (2005) Antibody arrays in cancer research. *Mol. Cell. Proteomics* **4**, 377–383
 50. Hudelist, G., Pacher-Zavisin, M., Singer, C. F., Holper, T., Kubista, E., Schreiber, M., Manavi, M., Bilban, M., and Czerwenka, K. (2004) Use of high-throughput protein array for profiling of differentially expressed proteins in normal and malignant breast tissue. *Breast Cancer Res. Treat.* **86**, 281–291
 51. Celis, J. E., Moreira, J. M. A., Cabezon, T., Gromov, P., Friis, E., Rank, F., and Gromova, I. (2005) Identification of extracellular and intracellular signaling components of the mammary adipose tissue and its interstitial fluid in high risk breast cancer patients: toward dissecting the molecular circuitry of epithelial-adipocyte stromal cell interactions. *Mol. Cell. Proteomics* **4**, 492–522
 52. Lin, Y., Huang, R., Chen, L., Li, S., Shi, Q., Jordan, C., and Huang, R. P. (2004) Identification of interleukin-8 as estrogen receptor-regulated factor involved in breast cancer invasion and angiogenesis by protein arrays. *Int. J. Cancer* **109**, 507–515
 53. Woodbury, R. L., Varnum, S. M., and Zangar, R. C. (2002) Elevated HGF levels in sera from breast cancer patients detected using a protein microarray ELISA. *J. Proteome Res.* **1**, 233–237
 54. Celis, J. E., Gromov, P., Cabezon, T., Moreira, J. M. A., Ambartsumian, N., Sandelin, K., Rank, F., and Gromova, I. (2004) Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. *Mol. Cell. Proteomics* **3**, 327–344
 55. Nielsen, U. B., Cardone, M. H., Sinskey, A. J., MacBeath, G., and Sorger, P. K. (2003) Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9330–9335
 56. Liotta, L. A., Espina, V., Mehta, A. I., Calvert, V., Rosenblatt, K., Geho, D., Munson, P. J., Young, L., Wulfschuh, J., and Petricoin, E. F., III (2003) Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* **3**, 317–325
 57. Cekaite, L., Haug, O., Myklebost, O., Aldrin, M., Ostenstad, B., Holden, M., Frigessi, A., Hovig, E., and Sioud, M. (2004) Analysis of the humoral immune response to immunoselected phage-displayed peptides by a microarray-based method. *Proteomics* **4**, 2572–2582
 58. Wang, X., Yu, J., Sreekumar, A., Varambally, S., Shen, R., Giacherio, D., Mehra, R., Montie, J. E., Pienta, K. J., Sanda, M. G., Kantoff, P. W., Rubin, M. A., Wei, J. T., Ghosh, D., and Chinnaiyan, A. M. (2005) Autoantibody signatures in prostate cancer. *N. Engl. J. Med.* **353**, 1224–1235
 59. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021
 60. Stults, J. T., and Arnott, D. (2005) Proteomics. *Methods Enzymol.* **402**, 245–289
 61. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**, 466–469
 62. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071–2077
 63. Somiari, R. I., Somiari, S., Russell, S., and Shriver, C. D. (2005) Proteomics of breast carcinoma. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **815**, 215–225
 64. Wulfschuh, J. D., McLean, K. C., Paweletz, C. P., Sgroi, D. C., Trock, B. J., Steeg, P. S., and Petricoin, E. F., III (2001) New approaches to proteomic analysis of breast cancer. *Proteomics* **1**, 1205–1215
 65. Wulfschuh, J. D., Sgroi, D. C., Kruttsch, H., McLean, K., McGarvey, K., Knowlton, M., Chen, S., Shu, H., Sahin, A., Kurek, R., Wallwiener, D., Merino, M. J., Petricoin, E. F., III, Zhao, Y., and Steeg, P. S. (2002) Proteomics of human breast ductal carcinoma in situ. *Cancer Res.* **62**, 6740–6749
 66. Vercoutter-Edouart, A. S., Lemoine, J., Le Bourhis, X., Louis, H., Boilly, B., Nurcombe, V., Revillion, F., Peyrat, J. P., and Hondermarck, H. (2001) Proteomic analysis reveals that 14-3-3 σ is down-regulated in human breast cancer cells. *Cancer Res.* **61**, 76–80
 67. Hondermarck, H., Vercoutter-Edouart, A. S., Revillion, F., Lemoine, J., el-Yazidi-Belkoura, I., Nurcombe, V., and Peyrat, J. P. (2001) Proteomics of breast cancer for marker discovery and signal pathway profiling. *Proteomics* **1**, 1216–1232
 68. Bini, L., Magi, B., Marzocchi, B., Arcuri, F., Tripodi, S., Cintonino, M., Sanchez, J. C., Frutiger, S., Hughes, G., Pallini, V., Hochstrasser, D. F., and Tosi, P. (1997) Protein expression profiles in human breast ductal carcinoma and histologically normal tissue. *Electrophoresis* **18**, 2832–2841
 69. Franzen, B., Linder, S., Alaiya, A. A., Eriksson, E., Uruy, K., Hirano, T., Okuzawa, K., and Auer, G. (1996) Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytochromes. *Br. J. Cancer* **74**, 1632–1638
 70. Franzen, B., Auer, G., Alaiya, A. A., Eriksson, E., Uruy, K., Hirano, T., Okuzawa, K., Kato, H., and Linder, S. (1996) Assessment of homogeneity in polypeptide expression in breast carcinomas shows widely variable expression in highly malignant tumors. *Int. J. Cancer* **69**, 408–414
 71. Bergman, A. C., Benjamin, T., Alaiya, A., Waltham, M., Sakaguchi, K., Franzen, B., Linder, S., Bergman, T., Auer, G., Appella, E., Wirth, P. J., and Jornvall, H. (2000) Identification of gel-separated tumor marker proteins by mass spectrometry. *Electrophoresis* **21**, 679–686
 72. Goufman, E. I., Moshkovskii, S. A., Tikhonova, O. V., Lokhov, P. G., Zgoda, V. G., Serebryakova, M. V., Toropygin, I. Y., Vlasova, M. A., Safarova, M. R., Makarov, O. V., and Archakov, A. I. (2006) Two-dimensional electrophoretic proteome study of serum thermostable fraction from patients with various tumor conditions. *Biochemistry (Mosc.)* **71**, 354–360
 73. Alexander, H., Stegner, A. L., Wagner-Mann, C., Du Bois, G. C., Alexander, S., and Sauter, E. R. (2004) Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin. Cancer Res.* **10**, 7500–7510
 74. Wall, D. B., Kachman, M. T., Gong, S., Hinderer, R., Parus, S., Misk, D. E., Hanash, S. M., and Lubman, D. M. (2000) Isoelectric focusing nonporous RP HPLC: a two-dimensional liquid-phase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. *Anal. Chem.* **72**, 1099–1111
 75. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999
 76. Pawlik, T., Hawke, D., Liu, Y., Krishnamurthy, S., Fritsche, H., Hunt, K., and Kuerer, H. (2006) Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. *BMC Cancer* **6**, 68
 77. Jessani, N., Niessen, S., Wei, B. Q., Nicolau, M., Humphrey, M., Ji, Y., Han, W., Noh, D. Y., Yates, J. R., III, Jeffrey, S. S., and Cravatt, B. F. (2005) A streamlined platform for high-content functional proteomics of primary human specimens. *Nat. Methods* **2**, 691–697
 78. Yanagisawa, K., Shyr, Y., Xu, B. J., Massion, P. P., Larsen, P. H., White, B. C., Roberts, J. R., Edgerton, M., Gonzalez, A., Nadaf, S., Moore, J. H., Caprioli, R. M., and Carbone, D. P. (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* **362**, 433–439
 79. Schwartz, S. A., Weil, R. J., Thompson, R. C., Shyr, Y., Moore, J. H., Toms, S. A., Johnson, M. D., and Caprioli, R. M. (2005) Proteomic-based prognosis of brain tumor patients using direct-tissue matrix-assisted laser desorption ionization mass spectrometry. *Cancer Res.* **65**, 7674–7681
 80. Chaurand, P., Sanders, M. E., Jensen, R. A., and Caprioli, R. M. (2004) Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. *Am. J. Pathol.* **165**, 1057–1068
 81. Issaq, H. J., Veenstra, T. D., Conrads, T. P., and Felschow, D. (2002) The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem. Biophys. Res. Commun.* **292**, 587–592
 82. Clarke, C. H., Buckley, J. A., and Fung, E. T. (2005) SELDI-TOF-MS proteomics of breast cancer. *Clin. Chem. Lab. Med.* **43**, 1314–1320
 83. Petricoin, E. F., and Liotta, L. A. (2004) SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr. Opin. Biotechnol.* **15**, 24–30

84. Kozak, K. R., Amneus, M. W., Pusey, S. M., Su, F., Luong, M. N., Luong, S. A., Reddy, S. T., and Farias-Eisner, R. (2003) Identification of biomarkers for ovarian cancer using strong anion-exchange ProteinChips: potential use in diagnosis and prognosis. *Proc. Natl. Acad. Sci. U. S. A* **100**, 12343–12348
85. Petricoin, E. F., III, Ornstein, D. K., Pawletz, C. P., Ardekani, A., Hackett, P. S., Hitt, B. A., Velasco, A., Trucco, C., Wiegand, L., Wood, K., Simone, C. B., Levine, P. J., Linehan, W. M., Emmert-Buck, M. R., Steinberg, S. M., Kohn, E. C., and Liotta, L. A. (2002) Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* **94**, 1576–1578
86. Zhang, Z., Bast, R. C., Jr., Yu, Y., Li, J., Sokoll, L. J., Rai, A. J., Rosenzweig, J. M., Cameron, B., Wang, Y. Y., Meng, X. Y., Berchuck, A., Van Haften-Day, C., Hacker, N. F., de Buijn, H. W., van der Zee, A. G., Jacobs, I. J., Fung, E. T., and Chan, D. W. (2004) Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res.* **64**, 5882–5890
87. Won, Y., Song, H. J., Kang, T. W., Kim, J. J., Han, B. D., and Lee, S. W. (2003) Pattern analysis of serum proteome distinguishes renal cell carcinoma from other urologic diseases and healthy persons. *Proteomics* **3**, 2310–2316
88. Koopmann, J., Zhang, Z., White, N., Rosenzweig, J., Fedarko, N., Jagannath, S., Canto, M. I., Yeo, C. J., Chan, D. W., and Goggins, M. (2004) Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. *Clin. Cancer Res.* **10**, 860–868
89. Diamandis, E. P. (2004) Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. *J. Natl. Cancer Inst.* **96**, 353–356
90. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P., Wilfond, B., Borg, A., and Trent, J. (2001) Gene-expression profiles in hereditary breast cancer. *N. Engl. J. Med.* **344**, 539–548
91. Palacios, J., Honrado, E., Osorio, A., Cazorla, A., Sarrío, D., Barroso, A., Rodriguez, S., Cigudosa, J. C., Diez, O., Alonso, C., Lerma, E., Dopazo, J., Rivas, C., and Benitez, J. (2005) Phenotypic characterization of BRCA1 and BRCA2 tumors based in a tissue microarray study with 37 immunohistochemical markers. *Breast Cancer Res. Treat.* **90**, 5–14
92. Jacquemier, J., Padovani, L., Rabayrol, L., Lakhani, S. R., Penault-Llorca, F., Denoux, Y., Fiche, M., Figueiro, P., Maisongrosse, V., Ledoussal, V., Martinez Penuela, J., Udvarhelyi, N., El Makdissi, G., Ginestier, C., Geneix, J., Charafe-Jauffret, E., Xerri, L., Eisinger, F., Birnbaum, D., and Sobol, H. (2005) Typical medullary breast carcinomas have a basal/myoepithelial phenotype. *J. Pathol.* **207**, 260–268
93. Moreira, J. M., Ohlsson, G., Rank, F. E., and Celis, J. E. (2005) Down-regulation of the tumor suppressor protein 14-3-3 σ is a sporadic event in cancer of the breast. *Mol. Cell. Proteomics* **4**, 555–569
94. Simon, R., Nocito, A., Hubscher, T., Bucher, C., Torhorst, J., Schraml, P., Bubendorf, L., Mihatsch, M. M., Moch, H., Wilber, K., Schotzau, A., Kononen, J., and Sauter, G. (2001) Patterns of her-2/neu amplification and overexpression in primary and metastatic breast cancer. *J. Natl. Cancer Inst.* **93**, 1141–1146
95. Van den Eynden, G. G., Van der Auwera, I., Van Laere, S., Colpaert, C. G., van Dam, P., Merajver, S., Kleer, C. G., Harris, A. L., Van Marck, E. A., Dirix, L. Y., and Vermeulen, P. B. (2004) Validation of a tissue microarray to study differential protein expression in inflammatory and non-inflammatory breast cancer. *Breast Cancer Res. Treat.* **85**, 13–22
96. Charafe-Jauffret, E., Tarpin, C., Bardou, V. J., Bertucci, F., Ginestier, C., Braud, A. C., Puig, B., Geneix, J., Hassoun, J., Birnbaum, D., Jacquemier, J., and Viens, P. (2004) Immunophenotypic analysis of inflammatory breast cancers: identification of an 'inflammatory signature'. *J. Pathol.* **202**, 265–273
97. Bertucci, F., Bernard, K., Llorid, B., Chang, Y. C., Granjeaud, S., Birnbaum, D., Nguyen, C., Peck, K., and Jordan, B. R. (1999) Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples. *Hum. Mol. Genet.* **8**, 1715–1722
98. Martin, K. J., Kritzman, B. M., Price, L. M., Koh, B., Kwan, C. P., Zhang, X., Mackay, A., O'Hare, M. J., Kaelin, C. M., Mutter, G. L., Pardee, A. B., and Sager, R. (2000) Linking gene expression patterns to therapeutic groups in breast cancer. *Cancer Res.* **60**, 2232–2238
99. Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O., and Botstein, D. (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9212–9217
100. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnson, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., and Botstein, D. (2000) Molecular portraits of human breast tumours. *Nature* **406**, 747–752
101. Korsching, E., Packeisen, J., Agelopoulos, K., Eisenacher, M., Voss, R., Isola, J., van Diest, P. J., Brandt, B., Boecker, W., and Buerger, H. (2002) Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. *Lab. Invest.* **82**, 1525–1533
102. Callagy, G., Cattaneo, E., Daigo, Y., Happerfield, L., Bobrow, L. G., Pharoah, P. D., and Caldas, C. (2003) Molecular classification of breast carcinomas using tissue microarrays. *Diagn. Mol. Pathol.* **12**, 27–34
103. Zhang, D. H., Salto-Tellez, M., Chiu, L. L., Shen, L., and Koay, E. S. (2003) Tissue microarray study for classification of breast tumors. *Life Sci.* **73**, 3189–3199
104. Jacquemier, J., Ginestier, C., Rougemont, J., Bardou, V. J., Charafe-Jauffret, E., Geneix, J., Adelaide, J., Koki, A., Houvenaeghel, G., Hassoun, J., Maraninchi, D., Viens, P., Birnbaum, D., and Bertucci, F. (2005) Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. *Cancer Res.* **65**, 767–779
105. Abd El-Rehim, D. M., Ball, G., Pinder, S. E., Rakha, E., Paish, C., Robertson, J. F., Macmillan, D., Blamey, R. W., and Ellis, I. O. (2005) High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int. J. Cancer* **116**, 340–350
106. Makretsov, N. A., Huntsman, D. G., Nielsen, T. O., Yorlida, E., Peacock, M., Cheang, M. C., Dunn, S. E., Hayes, M., Van De Rijn, M., Bajdik, C., and Gilks, C. B. (2004) Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically significant groups of breast carcinoma. *Clin. Cancer Res.* **10**, 6143–6151
107. Nielsen, T. O., Hsu, F. D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L., Akslen, L. A., Ragaz, J., Gown, A. M., Gilks, C. B., van de Rijn, M., and Perou, C. M. (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin. Cancer Res.* **10**, 5367–5374
108. Rakha, E. A., Putti, T. C., Abd El-Rehim, D. M., Paish, C., Green, A. R., Powe, D. G., Lee, A. H., Robertson, J. F., and Ellis, I. O. (2006) Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *J. Pathol.* **208**, 495–506
109. Ristimaki, A., Sivula, A., Lundin, J., Lundin, M., Salminen, T., Haglund, C., Joensuu, H., and Isola, J. (2002) Prognostic significance of elevated cyclooxygenase-2 expression in breast cancer. *Cancer Res.* **62**, 632–635
110. Poremba, C., Heine, B., Diallo, R., Heinecke, A., Wai, D., Schaefer, K. L., Braun, Y., Schuck, A., Lanvers, C., Bankfalvi, A., Kneif, S., Torhorst, J., Zuber, M., Kochli, O. R., Mross, F., Dieterich, H., Sauter, G., Stein, H., Fogt, F., and Boecker, W. (2002) Telomerase as a prognostic marker in breast cancer: high-throughput tissue microarray analysis of hTERT and hTR. *J. Pathol.* **198**, 181–189
111. Spizzo, G., Went, P., Dirmhofer, S., Obrist, P., Simon, R., Spichtin, H., Maurer, R., Metzger, U., von Castelberg, B., Bart, R., Stopatschinskaya, S., Kochli, O. R., Haas, P., Mross, F., Zuber, M., Dietrich, H., Bischoff, S., Mirlacher, M., Sauter, G., and Gastl, G. (2004) High Ep-CAM expression is associated with poor prognosis in node-positive breast cancer. *Breast Cancer Res. Treat.* **86**, 207–213
112. Brandt, B., Kemming, D., Packeisen, J., Simon, R., Helms, M., Feldmann, U., Matuschek, A., Kersting, C., Hinrichs, B., Bidart, J. M., Bellet, D., Bartkowiak, K., Dankbar, N., Dittmar, T., Sauter, G., Boecker, W., and Buerger, H. (2005) Expression of early placenta insulin-like growth factor in breast cancer cells provides an autocrine loop that predominantly enhances invasiveness and motility. *Endocr. Relat. Cancer* **12**, 823–837
113. Simon, R., Panussis, S., Maurer, R., Spichtin, H., Glatz, K., Tapia, C., Mirlacher, M., Rufe, A., Torhorst, J., and Sauter, G. (2004) KIT (CD117)-

- positive breast cancers are infrequent and lack KIT gene mutations. *Clin. Cancer Res.* **10**, 178–183
114. Ginestier, C., Bardou, V. J., Popovici, C., Charafe-Jauffret, E., Bertucci, F., Geneix, J., Adélaïde, J., Chaffanet, M., Hassoun, J., Viens, P., Jacquemier, J., and Birnbaum, D. (2003) Loss of FHIT protein expression is a marker of adverse evolution in good prognosis localized breast cancer. *Int. J. Cancer* **107**, 854–862
 115. Nevalainen, M. T., Xie, J., Torhorst, J., Bubendorf, L., Haas, P., Kononen, J., Sauter, G., and Rui, H. (2004) Signal transducer and activator of transcription-5 activation and breast cancer prognosis. *J. Clin. Oncol.* **22**, 2053–2060
 116. Callagy, G. M., Pharoah, P. D., Pinder, S. E., Hsu, F. D., Nielsen, T. O., Ragaz, J., Ellis, I. O., Huntsman, D., and Caldas, C. (2006) Bcl-2 is a prognostic marker in breast cancer independently of the Nottingham Prognostic Index. *Clin. Cancer Res.* **12**, 2468–2475
 117. Mehra, R., Varambally, S., Ding, L., Shen, R., Sabel, M. S., Ghosh, D., Chinnaiyan, A. M., and Kleer, C. G. (2005) Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. *Cancer Res.* **65**, 11259–11264
 118. Moyano, J. V., Evans, J. R., Chen, F., Lu, M., Werner, M. E., Yehiely, F., Diaz, L. K., Turbin, D., Karaca, G., Wiley, E., Nielsen, T. O., Perou, C. M., and Cryns, V. L. (2006) α B-crystallin is a novel oncoprotein that predicts poor clinical outcome in breast cancer. *J. Clin. Investig.* **116**, 261–270
 119. van de Rijn, M., Perou, C. M., Tibshirani, R., Haas, P., Kallioniemi, O., Kononen, J., Torhorst, J., Sauter, G., Zuber, M., Kochli, O. R., Mross, F., Dieterich, H., Seitz, R., Ross, D., Botstein, D., and Brown, P. (2002) Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. *Am. J. Pathol.* **161**, 1991–1996
 120. Stein, T., Price, K. N., Morris, J. S., Heath, V. J., Ferrier, R. K., Bell, A. K., Pringle, M. A., Villadsen, R., Petersen, O. W., Sauter, G., Bryson, G., Mallon, E. A., and Gusterson, B. A. (2005) Annexin A8 is up-regulated during mouse mammary gland involution and predicts poor survival in breast cancer. *Clin. Cancer Res.* **11**, 6872–6879
 121. Bertucci, F., Borie, N., Ginestier, C., Groulet, A., Charafe-Jauffret, E., Adélaïde, J., Geneix, J., Bachelart, L., Finetti, P., Koki, A., Hermitte, F., Hassoun, J., Debono, S., Viens, P., Fert, V., Jacquemier, J., and Birnbaum, D. (2004) Identification and validation of an ERBB2 gene expression signature in breast cancers. *Oncogene* **23**, 2564–2575
 122. Linke, S. P., Bremer, T. M., Herold, C. D., Sauter, G., and Diamond, C. (2006) A multimarker model to predict outcome in tamoxifen-treated breast cancer patients. *Clin. Cancer Res.* **12**, 1175–1183
 123. Li, J., Zhang, Z., Rosenzweig, J., Wang, Y. Y., and Chan, D. W. (2002) Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin. Chem.* **48**, 1296–1304
 124. Vlahou, A., Laronga, C., Wilson, L., Gregory, B., Fournier, K., McGaughey, D., Perry, R. R., Wright, G. L., Jr., and Semmes, O. J. (2003) A novel approach toward development of a rapid blood test for breast cancer. *Clin. Breast Cancer* **4**, 203–209
 125. Becker, S., Cazares, L. H., Watson, P., Lynch, H., Semmes, O. J., Drake, R. R., and Laronga, C. (2004) Surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) differentiation of serum protein profiles of BRCA-1 and sporadic breast cancer. *Ann. Surg. Oncol.* **11**, 907–914
 126. Hu, Y., Zhang, S., Yu, J., Liu, J., and Zheng, S. (2005) SELDI-TOF-MS: the proteomics and bioinformatics approaches in the diagnosis of breast cancer. *Breast* **14**, 250–255
 127. Mendrinos, S., Nolen, J. D., Styblo, T., Carlson, G., Pohl, J., Lewis, M., and Ritchie, J. (2005) Cytologic findings and protein expression profiles associated with ductal carcinoma of the breast in ductal lavage specimens using surface-enhanced laser desorption and ionization-time of flight mass spectrometry. *Cancer* **105**, 178–183
 128. Pawletz, C. P., Trock, B., Pennanen, M., Tsangaris, T., Magnant, C., Liotta, L. A., and Petricoin, E. F., III (2001) Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF: potential for new biomarkers to aid in the diagnosis of breast cancer. *Dis. Markers* **17**, 301–307
 129. Sauter, E. R., Zhu, W., Fan, X. J., Wassell, R. P., Chervoneva, I., and Du Bois, G. C. (2002) Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. *Br. J. Cancer* **86**, 1440–1443
 130. Fowler, L. J., Lovell, M. O., and Izbicka, E. (2004) Fine-needle aspiration in PreservCyt: a novel and reproducible method for possible ancillary proteomic pattern expression of breast neoplasms by SELDI-TOF. *Mod. Pathol.* **17**, 1012–1020
 131. Ricolleau, G., Charbonnel, C., Lode, L., Loussouarn, D., Joalland, M. P., Bogumil, R., Jourdain, S., Minvielle, S., Campone, M., Deporte-Fety, R., Campion, L., and Jezequel, P. (2006) Surface-enhanced laser desorption/ionization time of flight mass spectrometry protein profiling identifies ubiquitin and ferritin light chain as prognostic biomarkers in node-negative breast cancer tumors. *Proteomics* **6**, 1963–1975
 132. Goncalves, A., Esterni, B., Bertucci, F., Sauvan, R., Chabannon, C., Cubizolles, M., Bardou, V. J., Houvenaeghel, G., Jacquemier, J., Granjeaud, S., Meng, X. Y., Fung, E. T., Birnbaum, D., Maraninchi, D., Viens, P., and Borg, J. P. (2006) Postoperative serum proteomic profiles may predict metastatic relapse in high-risk primary breast cancer patients receiving adjuvant chemotherapy. *Oncogene* **25**, 981–989
 133. Heike, Y., Hosokawa, M., Osumi, S., Fujii, D., Aogi, K., Takigawa, N., Ida, M., Tajiri, H., Eguchi, K., Shiwa, M., Wakatabe, R., Arikuni, H., Takaue, Y., and Takashima, S. (2005) Identification of serum proteins related to adverse effects induced by docetaxel infusion from protein expression profiles of serum using SELDI ProteinChip system. *Anticancer Res.* **25**, 1197–1203
 134. Pusztai, L., Gregory, B. W., Baggerly, K. A., Peng, B., Koomen, J., Kuerer, H. M., Esteve, F. J., Symmans, W. F., Wagner, P., Hortobagyi, G. N., Laronga, C., Semmes, O. J., Wright, G. L., Jr., Drake, R. R., and Vlahou, A. (2004) Pharmacoproteomic analysis of prechemotherapy and postchemotherapy plasma samples from patients receiving neoadjuvant or adjuvant chemotherapy for breast carcinoma. *Cancer* **100**, 1814–1822
 135. Kuerer, H. M., Coombes, K. R., Chen, J. N., Xiao, L., Clarke, C., Fritsche, H., Krishnamurthy, S., Marcy, S., Hung, M. C., and Hunt, K. K. (2004) Association between ductal fluid proteomic expression profiles and the presence of lymph node metastases in women with breast cancer. *Surgery* **136**, 1061–1069
 136. Pawlik, T. M., Fritsche, H., Coombes, K. R., Xiao, L., Krishnamurthy, S., Hunt, K. K., Pusztai, L., Chen, J. N., Clarke, C. H., Arun, B., Hung, M. C., and Kuerer, H. M. (2005) Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry. *Breast Cancer Res. Treat.* **89**, 149–157
 137. Coombes, K. R., Fritsche, H. A., Jr., Clarke, C., Chen, J. N., Baggerly, K. A., Morris, J. S., Xiao, L. C., Hung, M. C., and Kuerer, H. M. (2003) Quality control and peak finding for proteomics data collected from nipple aspirate fluid by surface-enhanced laser desorption and ionization. *Clin. Chem.* **49**, 1615–1623
 138. Sauter, E. R., Shan, S., Hewett, J. E., Speckman, P., and Du Bois, G. C. (2005) Proteomic analysis of nipple aspirate fluid using SELDI-TOF-MS. *Int. J. Cancer* **114**, 791–796
 139. Bertucci, F., Finetti, P., Cervera, N., Charafe-Jauffret, E., Mamessier, E., Adélaïde, J., Debono, S., Houvenaeghel, G., Maraninchi, D., Viens, P., Charpin, C., Jacquemier, J., and Birnbaum, D. (2006) Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers. *Cancer Res.* **66**, 4636–4644
 140. Goldhirsch, A., Glick, J. H., Gelber, R. D., Coates, A. S., Thurlimann, B., Senn, H.-J., and Panel Members (2005) Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005. *Ann. Oncol.* **16**, 1569–1583
 141. Ermolaeva, O., Rastogi, M., Pruitt, K. D., Schuler, G. D., Bittner, M. L., Chen, Y., Simon, R., Meltzer, P., Trent, J. M., and Boguski, M. S. (1998) Data management and analysis for gene expression arrays. *Nat. Genet.* **20**, 19–23
 142. Heinrich, M. C., Corless, C. L., Demetri, G. D., Blanke, C. D., von Mehren, M., Joensuu, H., McGreevey, L. S., Chen, C. J., Van den Abbeele, A. D., Druker, B. J., Kiese, B., Eisenberg, B., Roberts, P. J., Singer, S., Fletcher, C. D., Silberman, S., Dimitrijevic, S., and Fletcher, J. A. (2003) Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. *J. Clin. Oncol.* **21**, 4342–4349
 143. Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S. R., Weiss, R. A., and Liotta, L. A. (1996) Laser capture microdissection. *Science* **274**, 998–1001