9 Proteomic Expression Profiling of Breast Cancer

Hans Neubauer, Tanja Fehm, Christina Schütz, Runa Speer, Erich Solomayer, André Schrattenholz, Michael A. Cahill, Raffael Kurek

Recent Results in Cancer Research, Vol. 176

Abstract

Breast cancer is one of the most common cancers observed in women in industrialized Western countries. The development of novel diagnostic methods and the application of modern systemic therapies have significantly optimized early detection and therapy of breast cancer. However, many patients are currently overtreated. Traditionally, tumours have been categorized on the basis of histopathological criteria. However, staining pattern and intensity of cancer cells are not sufficient to reflect the molecular events driving tumour development and progression. Therefore, new genomic, transcriptomic and proteomic techniques are applied to clinical samples aiming to identify new targets for a therapy tailored for an individual patient. After an introduction to common genomic and transcriptomic profiling technologies and their relevance for clinical use, we will focus on analytical and preanalytical applications for the identification of new therapeutic targets by protein profiling, with a special emphasis on two-dimensional gel-technologies (2D-PAGE), particularly as they apply to the study of breast cancer.

9.1 Introduction

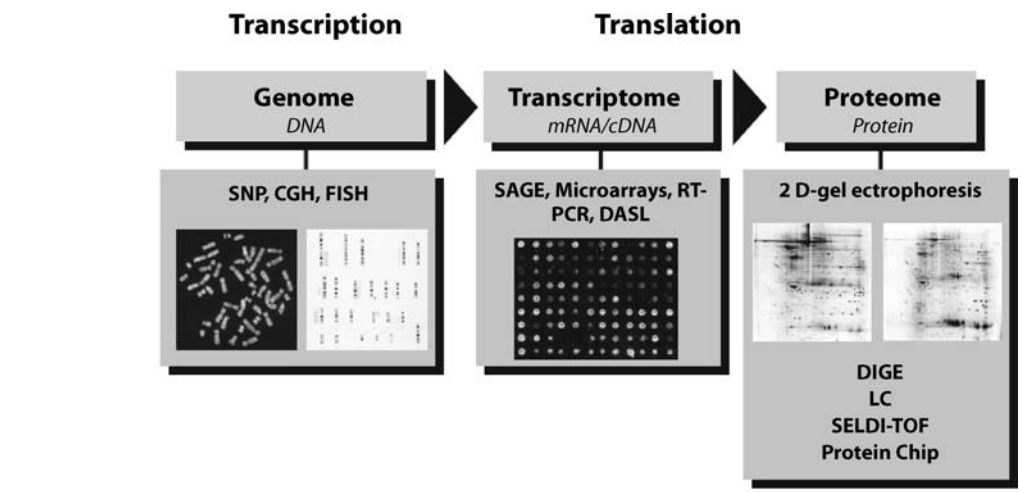
In the United States approximately 213,000 new cases of invasive breast cancer have been diagnosed in 2006, constituting approximately 31% of all new cancer cases among women; 41,000 disease-related deaths were expected (Jemal et al. 2006). In the past 10 years, improvements in diagnostic procedures for early detection and their broad application, together with the introduction of modern systemic therapies, have resulted in significant progress in early diagnosis and breast cancer therapy. In a new development, it has become feasible in clinical oncology to consider tailoring cancer therapy to an individual level of complexity by the use of suitable biomarkers. The current St. Gallen guidelines for the selection of adjuvant systemic therapy for early breast cancer patients include tumour size, grading, lymph node (LN) status, menopausal status, peritumoural vessel invasion, hormone receptor status and epidermal growth factor receptor 2 (HER2/neu) status. These markers can be classified into two major classes (Biomarkers Definitions Working Group 2001):

1. *Prognostic markers*. These provide information about the malignant potential of tumours, thereby predicting the outcome of a disease
2. *Predictive markers*. These are used to choose between different alternative treatment modalities. For instance, breast cancer patients with oestrogen receptor-positive tumours are usually treated with anti-oestrogen drugs such as tamoxifen and aromatase inhibitors, whereas oestrogen receptor-negative tumour patients are treated with chemotherapy. The overexpression of Her2/neu is predictive for the use of trastuzumab (Herceptin) at the same time as being a "drug target".

The use of these mainly histology-based prognostic parameters performs reasonably well using group-based statistical analyses. However, regarding outcome predictions for the individual

patient, these parameters should be supplemented with molecular parameters to reduce the uncomfortably high degree of uncertainty. Additionally, for prediction of therapy outcomes, the classical biomarkers provide only limited resolution of the manifested phenotype and have a limited capacity for individualizing a therapy. A more precise stratification of patients into responders versus non-responders to therapeutic agents is urgently needed by utilizing additional parameters. Therefore, new biological markers must be sought at all levels which are used to store the holistic biological information in a cell or tissue. These have been artificially classified into categories of convenience, such as the genome, the transcriptome, and the proteome and are dynamic, overlapping, and continuous in living systems (Fig. 9.1).

The introduction of microarray technology for nucleic acids opened the way to simultaneously analyse many genes—in contrast to classical histopathology—providing their specific profile of expression in a panoramic view. Therefore, a whole molecular profile is able to depict the polygenic origin of cancer, the multi-step process of tumourigenesis and the progression of cancer, which are reflected by genetic alterations that drive the transformation of normal human cells into highly malignant derivatives (Hanahan and Weinberg 2000). This concept of "molecular portraits" for each patient's tumour, first discussed by Perou et al., transcends histologic boundaries and indicates how array analysis can—compared to individual tumour markers—also provide new insights into breast cancer classification enabling a more refined stratification of the patients (Perou et al. 2000). Each of these molecular subtypes may be associated with a distinct clinical behaviour and treatment response as was shown for breast cancers of basal-like, HER2/neu over-



Modifications	Epigenetic e.g. Methylation	Genetic Point Mutation Gene Deletion Translocation Amplification	Posttranscriptional Splicing Editing	Posttranslational (>200) Splicing Phosphorylation Glycosylation etc.
Possible Amplification	Yes	Yes	Yes	No

Fig. 9.1 Different levels of biomarker research. Schematically depicted are the levels in a cell or tissue which are used to store biological information, possible modifications of genomic DNA, messenger RNA (mRNA) and proteins, and methods developed to identify them. DNA and mRNA can be amplified by polymerase chain reaction (PCR) and linear amplification methods, respectively

expressing, luminal A, and luminal B subtypes (Sorlie et al. 2001).

Besides the huge progress achieved in transcriptomics high-throughput analysis methods, recent technological progress has provided the tools to begin systematic development of comprehensive molecular pictures on the proteome level, investigating the component directly related to the phenotype and function of a cell (Hoheisel 2006).

9.2 Molecular Levels of Analytical Profiling

Since the publication of the draft sequence of the human genome (Lander et al. 2001; Venter et al. 2001), the aim of using information derived from genome analysis to tailor care to individual patients has gained prominence (Ginsburg and McCarthy 2001; Meyer and Ginsburg 2002; Snyderman and Williams 2003). Having access to the entire human genome sequence is a necessary prerequisite for molecular-based medicine. However, it is equally important to have the technology at hand to reliably visualize individual genomes, transcriptomes and proteomes providing information that, in combination with clinical data, can contribute to assessment of individual risks and guide clinical management and drug development.

9.2.1 Genomic Approaches

Genetic changes supporting oncogenesis can be point mutations, gene deletions ("loss of heterozygosity", LOH), translocations or amplifications. The exact number of genes in the entire human genome is currently estimated at about 35,000. Although humans, from a genetic point of view, are very similar to one another, there are single base exchanges in the DNA called single nucleotide polymorphisms (SNP) between individuals. It has recently been proposed that such SNPs arise by oxidation of genomic DNA, particularly involving 8-oxo-guanine, and that this is one of the largest sources of genomic diversity in human beings (Ohno et al. 2006). On average, one SNP is found every 1,500 base pairs, allowing the potential use of broad-based screens to pinpoint disease susceptibility genes to within a few 10,000ths of base pairs. Followed by sequencing of such short stretches, a specific genetic defect can be identified. Large-scale genotyping of SNPs is one new technology, assaying genotypes at thousands of loci (Hoheisel 2006). Other changes of the genome include cancer-specific variations in gene copy number, either through gene amplification or deletion. They initially highlighted the direct connection between such changes and disease, and are useful in diagnosis (Lichter et al. 1990; Feuk et al. 2006). One example in breast cancer is the amplification of the oncogene HER2/neu—methodologically determined on the genomic level by molecular methods such as polymerase chain reaction (PCR) and/or fluorescent in situ hybridization (FISH) (Benohr et al. 2005; Slamon et al. 1987; Sjogren et al. 1998). Another technology applied to analyse variations in gene copy number is comparative genomic hybridization (CGH), which has evolved from the standard hybridization of genomic DNA on metaphase spreads (Kallioniemi et al. 1992) to microarray-based CGH (array-CGH) (Pinkel et al. 1998). This technique facilitates the localization of copy-number changes very precisely by arraying probes derived from genomic sequences [e.g. BAC (bacterial artificial chromosomes) contiguous sequences (contigs)] that are tiled across a locus of interest. Such continuous coverage has been achieved for the entire human genome (Ishkanian et al. 2004; Hoheisel 2006). Beyond these genomic alterations, epigenomic modifications imposed onto the DNA, by e.g. environmental effects, can change gene expression and modify gene products in ways that initiate, accelerate or retard progression of pathologic processes without changing the coding nucleotide sequence of the genomic DNA (reviewed in: Laird 2005; Tlsty et al. 2004). The important implication is that in addition to the genetic analysis based upon nucleotide sequences, alternative approaches will be necessary to account for such environmental influences. A big practical advantage of genomic analysis is that the cellular component in question—the chromosomal DNA—is a fairly stable macromolecule enabling its convenient amplifi-

cation and analysis from formalin-fixed paraffin-embedded (FFPE) tissue archived in comprehensive tissue banks.

9.2.2 Transcriptomic Approaches

Analysing messenger RNA (mRNA) has always been an important and technology-enabled approach to examine the expression of genes. As an alternative to classical low-throughput Northern technologies, different high-throughput processes such as serial analysis of gene expression (SAGE) and microarray analysis have been developed to analyse global gene expression at the transcriptomic level (Velculescu et al. 2000; Schena et al. 1998). Both screening techniques provide the possibility to simultaneously evaluate the relative expression levels of large numbers of different mRNA transcripts in a panoramic view. Thereby, entire expression profiles themselves become a tumour marker that mirrors the polygenic nature of carcinogenesis. The increasing list of microarray experiments published each month reflects the straightforwardness of this technology, which is based on relatively inexpensive and easy-to-synthesize but sensitive and specific nucleic acid sequences. These can be arrayed into miniaturized standardized platforms and function as docking sites for complementary nucleotide sequences in the analyte (see several reviews: Butte 2002; Brentani et al. 2005; Hoheisel 2006). During the past decade, the microarray molecular profiling of breast tumours has produced a much more detailed classification scheme and has identified gene signature sets. One of the first attempts to characterize the variation in gene expression between sporadic breast tumour samples was published by Perou and co-workers (2000). In their groundbreaking study the authors showed that by differences in the expression profiles, breast cancer can be classified into so-called basal-like, HER2/neu overexpressing, luminal A, B, and C-, and normal-like tumours (Sorlie et al. 2001; Sorlie et al. 2003). These subtypes were correlated with overall survival and did not strongly reflect other clinical features such as LN status, tumour size or menopausal status, underscoring the importance of the molecular characterization of tumours. Another seminal DNA microarray analysis by van 't Veer and colleagues identified a 70-gene "profiler set" strongly predictive of a short interval to distant metastasis in LN-negative patients over 55 years of age with primary breast cancer and that can be used to classify primary breast carcinomas as having a gene-expression signature associated with either a poor or a good prognosis (van 't Veer et al. 2002; van de Vijver et al. 2002). Strikingly, this prognostic profiler set is independent of LN involvement, but is rather based upon its improved predictive power with respect to metastasis to non-lymphatic tissues. To prospectively evaluate this gene set which to date has only been retrospectively validated, a randomized clinical trial was launched in Europe ("MINDACT"). Prior to its launch the TRANSBIG Network embarked on an external, independent validation of the signature using frozen archival material of node-negative patients who are less than 60 years old. Preliminary analysis of approximately 300 samples from 6 different institutes shows that the overall performance of the 70-gene profiler set has a slightly reduced prognostic power in this external validation series compared to the original series published by van 't Veer et al. but it still outperformed the clinicopathologic risk assessment (Piccart et al. 2004; Buyse at al. 2006). Nevertheless, the "70-Gen-profiler set" from van 't Veer provides the platform for an already commercially available test. In addition to these studies further publications assigning the risk of patients with LN-negative breast cancer have been published recently (Wang et al. 2005; Pawitan et al. 2005). In one of these experiments, Wang et al. used oligonucleotide microarrays from Affymetrix to analyse a patient cohort with almost the same parameters as the cohort van 't Veer has used. They identified a 76-gene signature for untreated node-negative patients; it performed better in the multi-variate analysis compared with classical breast cancer prognostic factors. In summary, the results from van 't Veer et al. and Wang et al. indicate that gene-expression profiles are more powerful predictors of the outcome of disease in patients with breast cancer than the 2001 St. Gallen or the NIH consensus criteria and provide means for identification of patients needing adjuvant therapy (Eifel et al. 2001).

Other microarray studies have focussed on determining gene signatures of potential response of patients to specific chemotherapy and hormonal therapy regimens (Chang et al. 2003; Iwao-Koizumi et al. 2005; Ayers et al. 2004; Jansen et al. 2005). In the work by Chang et al. a discriminatory set of 92 genes was published which can be used to identify breast tumours responding to docetaxel. If the UICC (International Union Against Cancer) criteria had been applied to the same patients, some of the tumours in the gene expression-predicted resistant group would have belonged to the UICC-responding group, despite the fact that they had a completely different expression profile for the selected gene signature.

These exemplarily selected studies impressively show the power of microarrays in improving prognostic and predictive conclusions about breast cancer subtypes, but alternative test systems, which might be more suitable for routine use in the clinic, have been developed. For instance, to profile a small set of genes the application of quantitative RT-PCR (qRT-PCR) would be the most cost-effective and easy-to-handle method. This approach has been realized with the Oncotype DX test developed by Paik et al. and provided by the company Genomic Health (Paik et al. 2004). This test is based on real-time RT-PCR quantification of 16 cancer-related and 5 reference genes and is designed to identify patients benefiting most from adjuvant treatment with tamoxifen. Based on published data, 250 candidate genes were evaluated using three independent cohorts from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B.20 study to select the gene set mentioned above. After establishing a "recurrence score", the set was validated using a completely independent cohort from the NSABP B-14 trial. The assay allows for a better and/or more reproducible prognosis with oestrogen receptor (ER)-positive tumours in node-negative patients than the age of the patient, the size of the tumour or the histologic grade. Additionally, its quantitative read-out is also vastly superior to that of another commonly employed alternative method: immunohistochemistry.

An advantage of the RT-PCR approach is that it can be used with degraded RNAs derived from FFPE tumour samples, the most important and abundant source of clinical material. In contrast, microarray analysis is limited to frozen tumour samples imposing severe limitations on access to samples. Therefore, many laboratories are trying to develop methods that allow a similar degree of high-throughput gene-expression profiling using FFPE tissue as the starting material (Paik et al. 2005). This is not a trivial problem as RNAs extracted from FFPE tissues are chemically modified and fragmented and are therefore not ideal substrates for gene-expression profiling assays. Arcturus Biosciences (Mountain View, CA) has started marketing a reagent system called Paradise that has a combination of optimized RNA extraction and linear RNA amplification reagents. When used together with a specially designed oligonucleotide array, the GeneChip Human X3P array from Affymetrix (http://www.affymetrix.com/products/arrays/specific/x3p.affx), it promises to provide adequate gene-expression profiling data from FFPE-tissue. However, the "present call" rate for fresh paraffin blocks of breast cancer are reported to be below 30%, and it decreases to below 20% if the tissue block is more than 5 years old (Paik et al. 2005). Additionally, it has not been definitively clarified if gene sets obtained from cryo-preserved tissue correlate with gene sets from identical paraffin tissue (Sgroi et al. 2004; Ding et al. 2004).

Recently, a new method was developed—DASL (cDNA-mediated annealing, selection, extension and ligation) for gene-expression profiling to generate data from degraded RNAs such as those derived from FFPE tumour samples. This assay is a combination of microarray and qRT-PCR technologies into one platform that can be formatted to analyse the expression of a set of selected genes in a single clinical sample. It uses a minimal amount of total RNA (≤ 200 ng total RNA per assay) without prior linear amplification. The characteristics of DASL are (1) the use of FFPE samples as old as 24 years (unpublished data from Illumina); (2) high throughput with up to 96 clinical samples on one array plate and (3) the use of a custom gene panel with at least 512 genes per array (Bibikova et al. 2004; Fan et al. 2004).

The analysis of the transcriptome has already highlighted its potential for clinical usefulness by

generating promising results, but there are also big challenges waiting to be overcome. Especially regarding the use of microarrays, it is now evident that different platforms perform differently, resulting in only a marginal overlap of their gene-expression data, which is for instance the reason that to date no internationally accepted "risk-gene set" has been developed for "high-risk" breast tumour patients (Marshall 2004; Jenssen and Hovig 2005; Ein-Dor et al. 2005). This phenomenon is not restricted to breast cancer, but is also characteristic for gene sets of other diseases (Miklos and Maleszka 2004). Factors responsible for this phenomenon might include:

1. The application of different methodological standards
2. The application of different technical aspects including the type of array used (cDNA or oligonucleotide) (Marshall 2004), the gene sequences represented on the chips, the selection of tissue samples, the RNA extraction, what kind of probe was used for chip-hybridization (cDNA or cRNA), how the probe was labelled [cyanine (Cy) dyes, radioactivity, or biotin] and the conditions used for hybridization
3. The mode used for data processing, including the setting of instruments used for scanning the chips and the way the data were normalized and pre-processed
4. Differences in the study design and the number of samples included (Alizadeh et al. 2000)
5. Biological factors (Jarvinen et al. 2004; Ein-Dor et al. 2005)

9.2.3 Proteomic Approaches

The term "proteome" was coined in 1994 and is defined as the entire protein complement expressed by a cell line, tissue or organism. Proteomics, in analogy to genomics, is the study of the proteome, i.e. of all proteins—including their relative abundance, distribution, post-translational modifications, functions and interactions with other macromolecules—in a given cell or organism within a given environment at a specific stage in the cell cycle (Wasinger et al. 1995; Cai et al. 2004). Additionally, the study of the proteome also aims at the identification of protein isoforms. Proteomics has been classified into different sub-disciplines:

1. "Discovery-oriented" proteomics. Here, investigators are not able to impose their knowledge of biology on the experimental design. Such an experiment provides both known and unknown proteins.
2. "System-oriented" proteomics (MacBeath 2002; Choudhary and Grant 2004). In such an experiment a subset of proteins (e.g. a protein family) is directly analysed characterizing its biological functions, protein-protein or protein-DNA/RNA interactions, and protein post-translational modifications such as phosphorylation, sulphation or glycosylation.

Like the transcriptome, the rather dynamic nature of proteomes differs from individual to individual, and even from cell to cell. However, the sheer variety of potential modifications means that proteomic complexity dwarfs that of the transcriptome. Proteins undergo chemical modifications after they have been expressed, changing properties such as enzymatic activity, binding ability and activity. This myriad of modifications might give rise to 10–20 million chemically distinct polypeptides in a single tissue—a huge number compared to approximately 35,000 genes per cell of which around 6,000 are actively transcribed (Vuong et al. 2000). Despite this complicating aspect, the interest in applying proteomics to the identification of disease markers is increasing, because transcriptional activity does not necessarily reflect the activity of the proteins, which do all the work of the cell. When the same cells of tumours have been examined by both cDNA arrays and proteome methods, the correlation between mRNA transcript profiles and corresponding protein abundances has been reported to be only moderate (Anderson and Seilhamer 1997; Alaiya et al. 2000; Chen et al. 2003; Izzotti et al. 2004; Nishizuka et al. 2003). Therefore, it is obviously necessary to assess protein levels in instances where only protein expression levels correlate to disease. This is an important issue to consider, as most licensed tests that are available for disease detection are protein-based assays. In addition, proteomics provides the unique opportunity to develop serum markers to

be used for early disease detection and to follow treatment effects and disease progression.

The obvious advantages in analysing the cellular proteome, however, come with several complicating issues which are a consequence of its vast dynamic range of up to 10 orders of magnitude (Anderson and Anderson 2002), the plethora of post-translational modifications, boundless tissue, developmental and temporal specificities, disease and drug perturbations, and problems of sample degradation. Additionally, the chemistry of amino acids is much more complex to handle than that of nucleic acids: for proteins there is no amplification step that is analogous to the polymerase chain reaction (PCR). Hence, low-abundance proteins are often obscured by highly abundant proteins and separable protein species, e.g. cytoskeletal proteins, chaperones, endoplasmic reticulum proteins, proteasomal components and extra-cellular matrix proteins.

In order to get a deeper insight into the proteome of a cell, increasingly more sophisticated separation techniques have to be combined with highly sensitive mass spectrometry (MS) technologies for protein identification.

9.2.3.1 Mass Spectrometry

In recent years, MS has become almost a routine tool for identifying the proteins separated by different proteomic methods (Aebersold and Goodlett 2001; Domon and Aebersold 2006). Different types of mass spectrometers are used to support a range of research strategies in the protein sciences such as the determination of molecular weight, primary and higher order structure, post-translational modifications, quantitation, and localization. They differ in their physical principle, performance standards, mode of operation and ability to support specific analytical strategies. Matrix-assisted laser desorption/ionization (MALDI) ion sources are most commonly coupled with a time-of-flight (TOF) mass analyser. The method is very sensitive and quite tolerant to the presence of contaminants such as detergents or salts. However, these and other methods of ionization, separation and detection can be considered as modular units. Manufacturers are currently experimenting with various combinations of such modules to achieve powerful and sensitive separation and detection of peptide ions. There is an extensive variety of such combinations on the market, and it is difficult for the non-expert to maintain an overview, or for the expert to judge which system performs best. Therefore, this review aims to present an elementary level of introduction. In MALDI-TOF MS, peptides derived from proteolytic digested proteins are ionized from a plate into the spectrometer, and the mass to charge (m/z) ratios of peptides are measured based on the length of time for the peptides to move in a vacuum tube to reach a detector, and then a list of mass spectra is produced. Another method of protein ionization—electrospray ionization (ESI)—is most often coupled with ion-trap or triple quadrupole MS/MS spectrometer (Wilm et al. 1996). By applying database search algorithms, MS spectra are then matched to calculate masses in a sequence database, resulting in identification of target proteins, a method known as peptide mass fingerprinting (PMF; see Patterson and Aebersold 2003). Tandem mass spectrometry (MS/MS) can be thought of as a two-stage MS experiment whereby an ionized peptide is selected and allowed to pass into a collision cell for further fragmentation to determine sequential m/z values representing a series of ion fragments of the specific peptide (Arthur 2003).

The performance of TOF analysers has greatly improved in terms of resolution and accuracy, achieving mass accuracies in the low parts per million (ppm) range with appropriate internal calibration. However, for MS/MS measurements with TOF devices, such as the Bruker Ultraflex, calibration of the second MS step is not available, and the mass accuracy is up to 100 times poorer in MS/MS mode. While this is sufficient for confirmation of predicted fragment patterns from sequence databases to confirm peptide identity, it is frequently insufficient to permit de novo sequencing. The newer ion cyclotron resonance detectors offer the possibility of excellent mass accuracy for all measurements. These improvements offer levels of sensitivity and mass accuracy never before achieved for the detection, identification, and structural characterization of proteins. It is now possible to routinely measure molecular weights above 200 kDa as

well as obtain low parts per million mass measurement accuracy for the determination of peptides and proteins. Modern mass spectrometers can now rapidly map and fragment peptides that result from protease digestion to identify proteins and—supported by the rapid expansion of protein and gene databases—to obtain sequence information.

In addition to the MS applications combined with prior separation of proteins, imaging mass spectrometry (IMS) has been developed as a new technology enabling proteomic profiling direct on a tissue section (Chaurand et al. 2004). Briefly, molecules are desorbed from a sample that has been coated with an energy-absorbing matrix, which is a low molecular weight organic crystalline compound. The profiles recovered have been found to be extremely specific to a given tissue type and, when analysing serial sections, very reproducible. IMS offers the potential for the simultaneous analysis of many molecular species present in a single tumour regardless of the availability of specific antibodies or knowledge of the identity of the specific protein. However, its exact sensitivity is hard to estimate because exact amounts of proteins within a specific tissue are generally not well known. To date, profiling and imaging MS have been applied to multiple diseased tissues, including human non-small cell lung tumours, gliomas, and breast tumours (Yanagisawa et al. 2003; Schwartz et al. 2004; Chaurand et al. 2001) but this technique is still in its developmental phase.

One aspect of protein identification by MS that may not be apparent to more clinically oriented scientists is that the proteins are most often identified by comparing the pattern of ions measured in the mass spectrum with ion patterns predicted by comparison with sequence databases. This is true at the levels of measuring peptide masses (the process of peptide mass fingerprinting) as well as for measuring the size of ions generated when individual peptides are fragmented into smaller pieces. It is relatively infrequent that proteomics resorts to de novo interpretation of the amino acid sequence of an unknown protein without the assistance of sequence database information. This field has been reviewed recently (Domon and Aebersold 2006).

Traditionally, analysis of the proteins coded by genes was performed on single proteins at a time using techniques such as Western blots and immunoprecipitation. However, with the completion of the human genome project, proteomic technologies to identify and quantitate proteins on a global scale were developed, which are so far not as robust as those available for genomics and sometimes are still in their infancy and therefore constantly evolving. These technologies can be classified as gel-based and non-gel-based approaches and include two-dimensional (2D) gel electrophoresis-based methods including classical 2D-PAGE or 2D-difference gel electrophoresis (2D-DIGE); chromatographic separation techniques such as isotope coded affinity tag (ICAT) (Gygi et al. 1999b; Li et al. 2003) or multiple dimension protein identification technology (MudPIT) (Washburn et al. 2001; Chen et al. 2006); and recently the application of antibody and protein arrays (Table 9.1; Somiari et al. 2005).

9.2.3.2 2DE-Based Strategies

Despite several new technologies that have been introduced for high-throughput protein characterization and discovery, 2D-PAGE continues to be an affordable analytical methodology. Methods of 2D-PAGE can be divided into the conventional "one sample per gel" 2D techniques and the more recently developed DIGE (O'Farrell 1975; Patton 2002; Shaw et al. 2003; Somiari et al. 2005). They have recently been reviewed by Gorg et al. (2004), so we will limit our discussion to late technical developments and to applications to breast cancer. After visualizing by Coomassie blue, silver or fluorescent dye staining, each observed protein spot is quantified by its staining intensity. The major advantages of 2D-PAGE are the biochemical separation of intact polypeptide molecules, and their repertoire of post-translational modifications, including splicing variants. Any two species can be separated that differ in isoelectric point and molecular weight sufficiently to be separated with the resolution of the gel system used. Despite the utility of 2D-PAGE, its advantages are associated with several technology-related inherent disadvantages:

Table 9.1 Comparison of proteomic technologies and their contributions to biomarker discovery and early detection (Schrattenholz 2004; Wulfkuhle et al. 2003)

2D-PAGE[a]	2D-DIGE[b]+ radioisotopes[c]	2D-DIGE+Cy dyes[d]	LC[e]+/−s-isotopes[f]	SELDI[g]	Protein microarrays
Names of specific technologies with associated companies					
Various suppliers and companies	ProteoTope, ProteoSys	DIGE, Amersham Biosciences	ICAT[h], Applied Biosystems MudPIT[i]	SELDI Ciphergen	
Sensitivity					
Low (particularly for less-abundant proteins)	High (sub-attomole)	Low	Medium (~5 fmol)	Medium (femtomole) (diminishing yield at higher molecular weights)	Medium/high (picomole, depending on antibody)
Limited by detection methods (~1 fmol)					
Direct identification of markers					
Yes	Yes	Yes	Yes	No	Possible when coupled with MS technologies (Washburn 2003; Ouyang et al. 2003)
Pros					
Tried methodology	High resolution	Resolution	High throughput	Very high throughput	Flexible format
Good separation power	High dynamic range	No systemic error	Low molecular weight proteins	Low molecular weight proteins	Robust performance
	Differential quantification			Protein IDs not necessary for diagnostic pattern analysis	
	No systemic error				
	<1% crosstalk of labels				
Cons					
Low through put	Radioisotopes	Limited linear dynamic range of 2–3 orders of magnitude	Systemic error	Systemic error	Requires prior knowledge of analyte being measured
All IDs require validation and testing	Unlabelled protein for mass spectroscopy		False negatives	False negatives	Limited by sensitivity and specificity of antibody

Table 9.1 *(continued)*

2D-PAGE[a]	2D-DIGE[b]+radioisotopes[c]	2D-DIGE+Cy dyes[d]	LC[e]+/−s-isotopes[f]	SELDI[g]	Protein microarrays
Cons					
Time consuming			Limited dynamic range of 2–3 orders of magnitude	Very limited dynamic range	
Performs poorly for glycosylated proteins				Reproducibility issues need to be addressed	
Limited dynamic range (2–3 orders of magnitude)				Need for validation	
Use					
Discovery and identification of biomarkers	Discovery and identification of biomarkers	Discovery and identification of biomarkers	Discovery and identification of biomarkers	Diagnostic pattern analysis in body fluids and tissues	Multiparametric, systematic analysis of many analytes simultaneously
Throughput					
Low	Medium	Medium	High	Very high	High

[a] Two-dimensional polyacrylamide gel electrophoresis

[b] Two-dimensional difference gel electrophoresis

[c] Radioactive isotopes

[d] Cyanine dyes

[e] Liquid chromatography

[f] Stable isotopes

[g] Surface-enhanced laser desorption ionization

[h] Isotope-coded affinity tag

[i] Multi-dimensional protein identification technology

1. Standard protein staining in gels has moderate sensitivities (approx. 1 fmol) and clear disadvantages in terms of dynamic range of protein concentrations (Vuong et al. 2000). Neither conventional 2D gel electrophoresis methods nor any other proteome technology has the sensitivity of gene chip arrays (Gerling et al. 2003). The two to three orders of magnitude of linear dynamic range of 2D-procedures also do not match the protein reality, and thus a considerable amount of information is lost. Therefore, 2D-PAGE traditionally requires a large amount of protein starting material, and it can hardly be used to reliably detect and identify low-abundance proteins such as transcription factors. This leads to a bias in the presence of high-abundance "housekeeping" proteins in every protein database.
2. Some highly abundant proteins are not detectable by 2D-PAGE. Because of their high hydrophobicity they are not soluble in the detergent/urea buffer employed for isoelectric focussing (IEF) in the first dimension. Other proteins co-migrate with higher abundance proteins and are therefore not quantifiable by staining alone.
3. Applying 2D-PAGE is a time-consuming and labour intensive method.
4. Complex protein patterns of related samples, or even of multiple aliquots of the same sam-

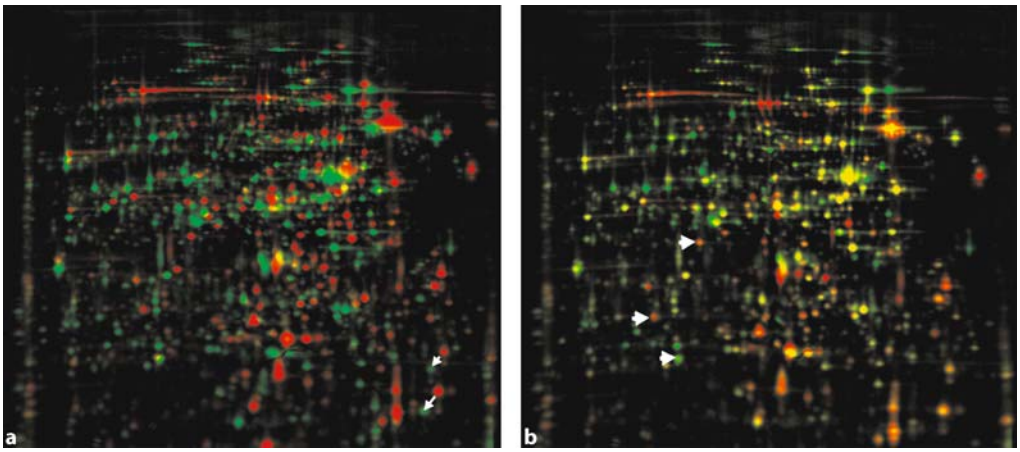


Fig. 9.2 a, b Differential analysis of 2D-PAGE. Depicted are overlays of two 2D-PAGE images in false colour. **a** Overlay of gels before warping. Protein spots of gel 1 are coloured in *green*, protein spots of gel 2 are in *red*. *Arrows* indicate orientation of warping. **b** Overlay of gels after warping. Protein spots equally represented on both gels result in a *yellow* colour. *Arrowhead* indicates a differentially displayed protein spot

ple, analysed in different gels are variable, i.e. matching of independent experiments is a severe problem. There are sophisticated software packages on the market such as Delta2D (Decodon, Greifswald, Germany) or ProteomeWeaver (Definiens) which permit gel-to-gel variations to be warped to each other. They are capable of quantifying the levels of proteins resolved on 2D-gels and have incorporated excellent spot detection algorithms and features that facilitate gel alignment and matching. Often included are commonly used bioinformatics tools such as principal component analysis, hierarchical clustering analysis or similar methods. Despite elaborate software, the inherent methodological variability necessitates some degree of manual workup for accurate spot matching and it binds a lot of computer power, making complex matching of independent experiments still quite laborious (Fig. 9.2).

5. One of the limitations of 2D-PAGE as a tool for biomarker discovery remains the sensitivity of MS characterization of protein spots, and particularly of post-translational modifications.
6. 2D-PAGE performs poorly in identifying heavily glycosylated proteins, since they tend to diffuse into clouds which are typically below the level of identification at any position in the gel.

The methodology of 2D-PAGE has also been improved by innovative modifications and sophisticated approaches: (1) the resolving power of the first dimension separation can be increased by the use of narrow pH-range immobilized pH gradient. A variation of this theme is the use of so-called "zoom gels" in which the protein contents of an individual sample are first fractionated into narrow pH ranges under low resolution, and then each fraction undergoes high-resolution separation by 2D-PAGE. Modern large "zoom gels" can reproducibly and reliably resolve thousands of different proteins from complex mixtures, which is superior to any other method (Gorg et al. 2004). Also, (2) sub-cellular fractionation of cells is able to increase the number of spots detected.

These different technological modifications together with advances in image analysis, data-mining and image storage have encouraged investigators to continue to apply 2D-PAGE for the analysis of complex samples leading to successful proteomic studies (Vercoutter-Edouart et al. 2001; Hondermarck et al. 2001; Bini et al. 1997;

Franzen et al. 1996a, b). These approaches have also been applied to biological fluids including serum (Goufman et al. 2006) and nipple aspirate fluids (NAF) (Alexander et al. 2004). In an exemplary study, Wulfkuhle et al. focussed on the identification of potential biomarkers in the early breast cancer lesion, ductal carcinoma in situ (DCIS), by analysing four cases of patient-matched, normal ductal epithelial cells and DCIS cells, specifically isolated from primary tissue by microdissection (Wulfkuhle et al. 2002, 2003). The proteomic profiles were compared by 2D-PAGE, differentially expressed spots were selected and sequenced by MS. The differential expression pattern for a subset of the identified proteins was validated by immunohistochemistry with a small, independent cohort of patient-matched normal/DCIS specimens. Very recently an analysis that combined 2D-PAGE with silver staining and MALDI-TOF and/or immunoblotting in sets of microdissected malignant breast epithelium and corresponding adjacent normal breast epithelia from 5 patients with invasive breast carcinoma was published (Hudelist et al. 2006).

9.2.3.3 2D-Difference Gel Electrophoresis

DIGE is a fairly recent advancement of 2D-PAGE technology, improving sample throughput and greatly enhancing gel reproducibility (Tonge et al. 2001; Von Eggeling et al. 2001). By using this method, which analyses several protein samples in one experiment, i.e. in one 2D-gel the systematic error of variable gel images can be avoided. Protein samples are labelled prior to electrophoresis with spectrally resolvable fluorescent cyanine dyes (Cy2, Cy3, and Cy5), Alexa dye, or with radioactive isotopes ("ProteoTope", ProteoSys, Mainz, Germany). The samples are then mixed prior to IEF and resolved on the same 2D-PAGE. The gel is scanned measuring the different labels, and multiple images corresponding to different samples are generated. Sample multiplexing in DIGE greatly refines the detection of changes at the protein level between samples (Tonge et al. 2001). Variation in spot intensities due to experimental factors, for example protein loss during sample entry into the strip, will be the same for each sample within a single DIGE gel. Therefore, the relative amounts of a protein between samples in a gel will be unchanged, thus increasing the confidence with which protein differences can be both detected and quantified. It reduces the amount of experimental variation due to a combination of multiple sample analysis in a single gel and internal standard correction (Alban et al. 2003; Lilley and Friedman 2004; Chen et al. 2005).

The fluorophores used for labelling are structurally similar and undergo nucleophilic substitution with the ε-amino group of lysine residues forming an amide. They have very similar molecular masses and are positively charged to match the charge that is replaced on the lysine residue. This matching of charge and mass ensures that all the samples essentially co-migrate to the same point during electrophoresis. In the labelling reaction, the dye/protein ratio is low. This ensures that protein molecules are only labelled with a single dye molecule. Quantitative cysteine alkylation such as DIGE labelling requires the correct stoichiometry of cysteine-reactive reagent, as well as correct reaction conditions to avoid generating artefactual spots caused by under- or over-alkylation, and these effects become critical for limited substrate levels in small volumes (Cahill et al. 2003; Sitek et al. 2005).

The fluorescence conventional strategy, with sensitivities similar to silver staining yet producing much better quantitative data, provides a linear dynamic range of almost three orders of magnitude; however, this range is reduced under conditions of signal multiplexing, so that direct multiplexing within one gel is not advisable with DIGE saturation labelling, due to fluorescence resonance energy transfer between the CyDyes (Patton 2002; Gruber et al. 2000). Accordingly, the considerable improvement in estimation precision for the ratio of abundance that can be ideally achieved by measuring intra-gel multiplexed protein spots cannot be realized with CyDyes (Poznanovic et al. 2005b). Recently the so-called saturation labelling method was introduced, which is now gaining in acceptance, and it has been applied to microdissected breast cancer samples (Wilson et al. 2005). Therein, one dye provides a reference standard to align spot patterns from multiple replicate gels (Kondo et al. 2003; Sitek et al. 2005). The disadvantage is that to reliably obtain statistically robust differential

expression data between microdissected samples sufficient multiple gels, with associated inter-gel variability, are required with micrograms of protein.

An alternative DIGE-method is radioactive labelling of the proteins using ProteoTope. Proteins from two samples are iodinated under chemically identical conditions with either 125I or 131I, mixed, co-electrophoresed by 2D-PAGE, and the signals from each isotope are differentially detected by ProteoTope imaging. To optimize the differential quantification of radioactive 2D-PAGE protein spots and to decrease the overlap in signals detected from different spots, high-resolution, 54-cm, immobilized pH gradients (IPGs)—IEF either in the continuous 54-cm IPG format or as serially connected 3×18 cm daisy chain IPG format—was developed (Poland et al. 2003; Poznanovic et al. 2005a). ProteoTope was intentionally designed for optimal protein spot quantification by radioactive detection and to achieve the most efficient analytical application of protein samples from extremely small sources, such as dissected tissue samples (Neubauer et al. 2006). It can reach a sensitivity of sub-attomole levels, with a dynamic range of over six orders, which is at least two orders of magnitude superior to fluorescent techniques. The linear dynamic range of detection of individual multiplexed, differentially abundant protein spots is typically greater than 15,000-fold, and cross-talk between the signals measured from the two samples is less than 1%. Under typical conditions, the labelling stoichiometry iodinates only approximately one tyrosine per 4,000 kDa, which generates sufficient radioactive signal, and provides a potential margin of error to establish reaction labelling conditions to avoid over-labelling of extremely limited samples. With ProteoTope, a statistically significant quantification of changes in the 15%–20% range can be distinguished with high certainty (Cahill et al. 2003), enabling quantification of even subtle protein changes in kinetic experiments. This greatly increases the accuracy of estimation of abundance ratios of identical proteins from different samples without the notorious complication of the inter-gel variability associated with conventional 2D-PAGE. Using this highly sensitive method also enables crosslabelling experiments which are performed for all labelling reactions to exclude the detection of false-positive proteins obtained by labelling or processing artefacts (Fig. 9.3).

ProteoTope was applied to a set of microdissected invasive breast cancer samples, all of which were ER+. Sub-pools were compared that were either positive or negative for the progesterone receptor (ER+/PR+ versus ER+/PR-). Employing a sample pooling strategy, several proteins differentially abundant, depending upon the presence or absence of PR, were found. In this experiment, approximately 180 ng of labelled protein was loaded per 54-cm analytical gel (Neubauer et al. 2006).

One issue using radiolabelled proteins is that the protein spots in analytical gels typically cannot be used for MS because of the vanishingly small amounts loaded. Therefore, preparative 2D-tracer-gels have to be run with high microgram amounts of unlabelled protein spiked with radioactive analytical protein, consuming highly valuable cryo-conserved tumour tissue. In the experiment mentioned above, 240 µg protein was required for the preparative tracer-gel (Neubauer et al. 2006).

9.2.3.4 Proteome Platforms Not Involving 2DE

As an alternative to gel-based separation approaches, attempts are being made to develop separation methods not involving 2D-PAGE. Liquid chromatography (LC) is one such technology that can successfully resolve a mixture of proteins and allow the isolation of individual proteins based on biochemical property. Some commonly used LC columns include size exclusion LC, ion exchange LC and reverse-phase LC. 2D LC, which combines pH gradient and reverse-phase columns, is a new proteomics technique that promises to extend the range of protein separation. Alternative "non-gel-based" approaches, such as MudPIT, have already been used effectively to catalogue many polypeptides in total protein mixtures from several organisms (Koller et al. 2002; Whitelegge 2002). However, while MudPIT is an excellent means of generating an exhaustive catalogue of proteins present in a particular protein sample, it does not yield reproducible quantitative information (Rose et al. 2004). MudPIT involves tryptic digestion of protein mixture followed by multi-dimensional

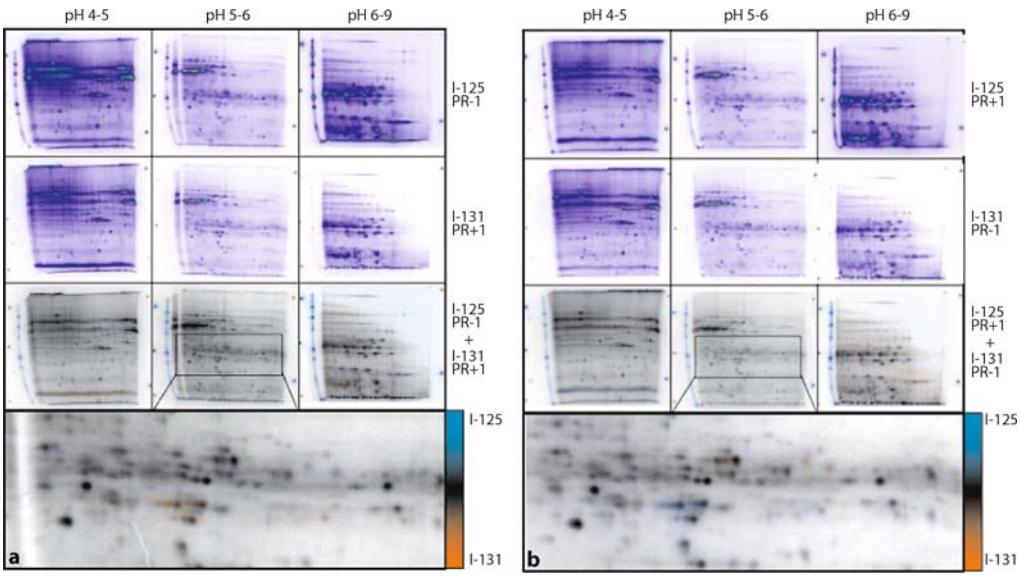


Fig. 9.3 a, b A 54-cm daisy chain IEF differential ProteoTope analysis of pooled LCM samples. The panels show actual images from an inverse replicate labelled ProteoTope experiment for one sample pair. **a** Analysis of sample PR-1 labelled with 125I, differentially compared with sample PR+1labelled with 131I. The *upper panels* show the signal detected for each isotope, depicted in false spectral colour. The signals for each isotope have been normalized against each other for total relative intensity in the *lower* dual channel images, where the signal for 125I is *blue*, the signal for 131I is *orange*, and equal amounts of both signals produces *grey* or *black* signal. Two pure sources each of 131I and 125I, as well as a 50% mixture of both isotopes, are measured on round, 2-mm pieces of filter paper placed next to each gel as imaging controls. Cross-talk between the signals from each isotope is <1%. The pH ranges of the 18-cm IPGs used for serial IEF are indicated *above* the panels, and the radioactive iodine isotope signals depicted in each panel are indicated on the *right*. Approximately 180-ng protein from each sample was loaded to each gel, and the above result was obtained by labelling approximately 3.6 mg protein from each sample. **b** The *top panels* show the inverse replicate experiment of **a**, where sample PR-1 is labelled with 131I, and sample PR+1 is labelled with 125I. (Reproduced with publisher's permission from Neubauer et al. 2006)

LC separating proteins by size-exclusion or cation exchange chromatography and in the second dimension by reverse-phase HPLC (Wall et al. 2000). Followed by MS measurement and database searching, the fractionated proteins can be directly analysed. This method shows advantages over gel-based techniques in speed, sensitivity, scope of analysis and dynamic range and it could be amenable to automation. Unfortunately, LC/MS is not capable of determining protein abundance. MudPIT has recently been associated with enzyme activity profiling in human tumour tissues, including breast tumours, and has generated functional signatures that correlate with previously described molecular subtypes (Jessani et al. 2005).

For quantitative comparisons of proteomes without the use of 2D-PAGE, ICAT technology has been developed recently (Gygi et al. 1999a). It is based on labelling a pair of samples simultaneously at the cysteine residues with differentially deuterated d0- and then d8-ICAT reagents ($^{13}C/^{12}C$ pairs are also available) (Gygi et al. 1999a). Unlike 2D-PAGE and SELDI-TOF (discussed later), which comparatively profile the naturally occurring forms of peptides and proteins, ICAT analyses the relative amounts of cysteine-containing peptides derived from e.g. tryptic digestion of protein extracts. The samples are then combined and analysed by LC-MS/MS. Each cysteinyl peptide appears as a pair of signals differing by the mass differentially encoded in the mass tag. The ratio of these signal intensities precisely indicates the ratio of abundance of the protein from which the peptide originates and the MS/MS spectrum of the peptide allows

the protein to be identified. ICAT has greatly expanded the range of proteins that can be analysed, quantified, and identified using these techniques. It has been applied to compare NAF from tumour-bearing and contralateral disease-free breasts of patients with unilateral early-stage breast cancer (EBC), identifying and quantifying differences in various specific protein expressions (Pawlik et al. 2006).

LC methods resolve hundreds to many thousands of peaks, and the limitations to the method involve not only generating reproducible analysis conditions, but also being able to process the vast reams of data that can be generated by e.g. MudPIT. The use of mass spectrometers with high-resolution Fourier transform ion cyclotron detectors reduces the stringency required of pre-mass spectrometric biochemical separations, and greatly increases the confidence with the protein identifications that are obtained (Haas et al. 2006). LC or chromatographic arrays have moderate sensitivities (~5 fmol) and a clear disadvantage in terms of dynamic range of protein concentrations—with moderate dynamic ranges of approximately three orders of magnitude (Schrattenholz 2004).

In a recently published study by Komatsu et al., 2D-LC and 2D-DIGE were compared with 2D-PAGE in combination with Coomassie brilliant blue (CBB) staining for their ability to identify proteins regulated by gibberellin (GA) in rice (Komatsu et al. 2006). Using 2D-LC and 2D-DIGE, many more proteins were detected compared with 2D-PAGE followed by CBB staining. Additionally, the two former methods detected proteins that were not reported previously. The difference between 2D-DIGE and 2D-LC was that minor GA-responsive proteins were detected only by 2D-DIGE and the low-molecular-weight proteins were detected only by the 2D-LC system. This suggests that the 2D-LC technique is the preferred method for detecting low-molecular-weight proteins. However, poor reproducibility and the large number of replicates required to establish statistical significance are problems that still must be resolved. Further, the 2D-DIGE technique is more sensitive and is able to make exact quantitative comparisons.

An often-cited technical disadvantage of "shotgun" proteomics methods such as MudPIT and ICAT, which reduce polypeptides to peptides before separation, is that modifications such as splice variants and post-translational protein modifications often escape detection. Additionally, although these methods are frequently designated as being high-throughput, they are typically expensive in terms of machine time, and experimental repetitions to gain meaningful sample sizes for the assessment of the statistical significance of differences are rarely performed for reasons of sample availability, analysis time and running costs. Nevertheless, these methods do generate considerable amounts of data in a short time.

9.2.3.5 Protein Chips

The successful application of DNA microarrays to genome and transcriptome research demonstrated the value of array-based measurements, and it was soon recognized that the ability to perform such experiments to measure proteins likewise would be very valuable. However, protein microarray technology is not as straightforward as DNA-based microarrays owing to the complex structure of proteins and e.g. the absence of a protein amplification method (Haab et al. 2001; MacBeath and Schreiber 2000). Protein microarrays are generally of two types (Speer et al. 2005): (1) antibody arrays, also known as forward phase arrays (FPA), in which the bait molecule—typically an antibody—is spotted onto suitable surfaces, and bound antigens are detected using radioactivity, fluorescence and chemiluminescence (Haab 2005). In FPAs, each spot represents only one type of bait molecule. The array is incubated with only one test sample that contains several different analytes of interest. The captured analytes are detected with a second tagged molecule or by labelling the analyte directly; and (2) non-antibody or reverse phase arrays (RPA) where sets of labelled proteins or even entire proteomes are spotted onto a slide (Liotta et al. 2003; Speer et al. 2005; Sheehan et al. 2005) (Fig. 9.4). The RPA design enables high-throughput analysis in the sense that every single spot comprises the entire protein pool of one patient/sample. Dozens of samples can be spotted on one slide in parallel and therefore can be probed with an antibody against the endpoint of interest in one experiment. This leads to excellent comparability

among samples and reduces errors. In contrast to FPAs, where different antibodies with various binding characteristics and capacities need to be combined on one platform, the RPA enables the user to choose the optimum binding conditions and dilution for the chosen antibody in order to detect the antigen of interest. This format allows multiple samples to be analysed under the same experimental conditions for any given analyte. As little as 30–60 µl of cell lysate is sufficient to print 50 or more arrays (Espina et al. 2003). Recently, a two-colour comparative fluorescence strategy has been used to compare protein levels between malignant and normal breast tissues from the same patient (Hudelist et al. 2004). Therein, a reference sample was 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.

The application of protein microarrays provides unique advantages. It offers high-throughput capabilities because of the low sample volumes needed. Protein chips also offer high robustness, sensitivity, inter-sample comparability and the possibility of a quantitative analysis of protein expression. Depending on the individual affinities of the immobilized antibodies, antigens can typically be detected in picomole (pmol) amounts, which is currently more than 1,000

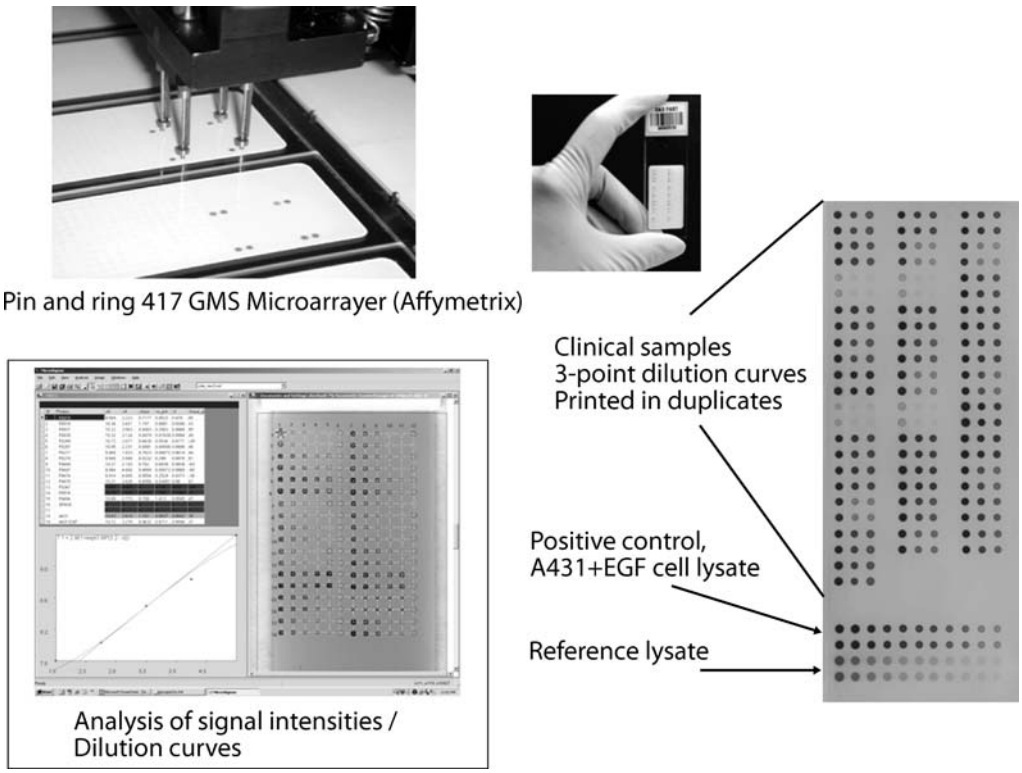


Fig. 9.4 Gene expression profiling by reverse-phase protein microarrays. The cell lysates are printed in duplicates at distinct positions on nitrocellulose coated slides using, for instance, a pin and ring 417 GMS Microarrayer (Affymetrix). Samples are arrayed in three-point dilution curves. The slides are then incubated with an antibody against a target protein of interest, such as phosphorylated endoplasmic reticulum kinase in this case. The antibody is detected by chemiluminescent, fluorescent or colourimetric assays. The intensity of the signal is proportional to the concentration of the target protein. A positive control lysate (A431 squamous carcinoma cell line) is printed on the array for monitoring immunostaining performance. Phosphorylation-specific reference peptides are printed in a 12-point dilution curve on the bottom of the array for comparative, precise quantification of patient samples between arrays. Image analysis is performed with, e.g., Microvigene (VigeneTech). *EGF*, epidermal growth factor

times less sensitive than 2D-PAGE or MS-based methods. However, with the development of suitable high-affinity antibodies and sensitive detection techniques, improvements are possible. For example, considering the specificity, affinity and cross-reactivity, only about 5% (MacBeath 2002) to 30% (Haab et al. 2001) of commercial antibodies are suitable for microarray-based analyses. Furthermore, as in any format of antibody array used, the specificity of antibody binding should be characterized, and the binding levels observed by microarray should be validated by independent methods. Appropriate assays like Western blots work well to confirm binding to a single target in a complex mixture, or to confirm changes in the level of certain targets (Orchekowski et al. 2005; Sreekumar et al. 2001). A method of validation that does not require purified antigen is immunoprecipitation and MS analysis of the captured proteins. This procedure should reveal both the specific and non-specific proteins bound by a particular target. The observation of the same biological information from different antibodies that bind to different epitopes on a protein can be a strong confirmation of the validity of the result.

Another characteristic of reverse phase protein microarrays is their flexibility, considering the fact that the design enables the user to analyse protein samples in denatured and non-denatured conditions. This flexibility ensures that protein microarrays can be used for a plethora of applications, such as drug discovery, biomarker identification, molecular profiling, developing "circuit maps" of on-going signal transduction in cell and tissue samples and assessment of response profiles for new drugs. Protein microarrays also provide a well-controlled in vitro way to study protein function, including protein–protein, protein–lipid and protein–nucleic acid interactions on a genome-wide basis (Jessani et al. 2002; Bulyk et al. 1999; Zhu et al. 2000). Already antibody arrays have been used to measure phosphorylation states and to study signalling in networks of interacting proteins (Gembitsky et al. 2004; Nielsen et al. 2003). Another novel use of antibody microarrays is to profile enzyme activity in complex proteasomes (Sieber et al. 2004). The measurement of enzyme activity rather than just abundance is important for determining the functional state of certain proteins and may be valuable for cancer research. Complex protein samples are treated with fluorescent activity-based probes and the labelled enzymes are captured and detected on antibody microarrays targeting those enzymes. Recently, expression profiling using reverse-phase protein microarrays has been applied investigating breast cancer tissue (Cowherd et al. 2004; Hudelist et al. 2004).

9.2.3.6 Surface-Enhanced Laser Desorption Ionization

The ProteinChip System (Ciphergen, Fremont, CA) is an alternative array system which has the advantage of its direct integration with MS. It is a rapid and sensitive analytical method which allows the quantification of proteins with different masses originating from complex protein mixtures such as body fluids or cell and/or tissue extracts by surface-enhanced laser desorption ionization (SELDI) at the femtomole level (He and Chiu 2003). However, SELDI cannot directly determine the identity of proteins. The SELDI protein chip platform is based on the principle that proteins from crude mixtures are selectively attracted to specific biochemical surfaces. Potential biomarkers may show a higher binding affinity to certain surfaces than serum albumin, haptoglobin and other abundant serum proteins. The current platform, based on nine different capture agents per chip, is claimed to achieve comprehensive coverage of the proteome. The SELDI-protein chip surfaces are chemically or biochemically modified to enable the capture of a certain group of proteins: they include (1) weak cation exchange, (2) strong anion exchange, (3) immobilized metal affinity chromatography, (4) reverse phase and normal phase and (5) biochemical affinities allowing proteins or antibodies to be bound directly to the chip (Arthur 2003). The retained proteins are then ionized and analysed using TOF/MS. The great advantages of this approach are its sensitivity to analyse small amounts of raw protein samples and its ability to detect proteins with molecular weights lower than 6 kDa. These characteristics make SELDI ProteinChip technology very attractive for biomarker discovery, especially because new biomarkers should be able to detect early forms of

cancer before it has metastasized throughout the body and should monitor a patient's response to therapy or the potential of recurrence in real time. Beyond that, for routine clinical use a reliable minimally invasive diagnostic test would be highly preferable. Serum has the advantage of being a readily accessible body fluid that is protein-rich and that is well-suited to proteomic analysis. Differences in proteomic patterns in serum from cancer patients can be due to: (1) serum proteins that are differentially expressed in patients with cancer, or serum proteins that are cleaved or modified in cancer patients; (2) proteins that are secreted by tumour cells; or (3) intracellular tumour proteins that are released when tumour cells die. Molecules released by tumour cells subsequently enter the blood and/or other fluids. Analysis of serum is pretty straightforward: it requires minimal sample preparation and only a very small amount of sample (1–20 μl). As an approach to serum biomarker discovery, proteomic pattern analysis has been developed to identify novel markers by comparing samples from patients with disease with those from healthy subjects. It is based on the analysis of large amounts of mass spectrometric data derived from complex protein mixture and does not per se require that the proteins involved are identified. Differentially displayed MS signals can then be identified and confirmed as potential biomarkers. In breast cancer, SELDI-TOF MS was used to investigate serum/plasma (Li et al. 2002; Vlahou et al. 2003; Becker et al. 2004; Hu et al. 2005), but also NAF (Mendrinos et al. 2005) and tumour tissues (Ricolleau et al. 2006) as a potential source for diagnostic biomarkers, and as a potential tool to predict outcome (Heike et al. 2005). Recently, SELDI has been applied to examine proteomic changes that occur in response to paclitaxel chemotherapy or 5-fluorouracil, doxorubicin and cyclophosphamide chemotherapy in plasma of 69 breast cancer patients and normal volunteers (Pusztai et al. 2004).

As with many of the initial SELDI-studies, standardization and independent validation using larger numbers of specimens is required to ensure the performance of these selected biomarkers. Attempts to validate serum proteome patterns are currently being made by collaborating research groups, where each sample is analysed by all participating laboratory. Encouraging results have been reported in reproducing separation patterns and disease classification in different laboratories (Grizzle et al. 2003–2004). Even though these results are promising, SELDI-TOF-MS screening for serum biomarker discovery limited the performance of this approach for biological fluids because proteomic technologies based on MS may not be sensitive enough to detect low-abundance molecules that are released by a few tumour cells or their microenvironment into the circulation. Profiling tissue extracts might therefore be more suitable for such a biomarker screening technology, as dilution effects are reduced (Ricolleau et al. 2006). Additionally, the clinician should be aware that doubts have been expressed concerning the reproducibility of the SELDI methodology itself (Diamandis 2004; Baggerly et al. 2004; Bons et al. 2005). Major issues to be solved are variations in serum collection and sample handling and how they will affect the analyses. Therefore, initiatives for standardizing the pre-analytical methods on different operational levels might help us to exploit the full power of the increasingly powerful high-throughput analytical technologies (Carr et al. 2004; Bradshaw et al. 2005).

9.3 Standardization of Pre-analytical Methods

The growing application of molecular diagnostic techniques in the clinic and in translational oncologic research has made it necessary to consider standardization of the pre-analytical methods used to collect, store and catalogue valuable human tissue. Therefore, guidelines for "banking" fresh tumour and normal tissues as a part of the routine activity in surgical pathological laboratories, specific for the requirements of molecular-based tests and quality control measures, have to be established. Additionally, various data about the tissue, the patient and the tissue inventory have to be managed and made available to the different co-operating disciplines and departments at different locations. For molecular analysis in the laboratory, guidelines for specimen handling, tissue enrichment strategies and quality controls must be established that are appropriate for the requirements of molecular-based tests. Finally, new experimental strategies

might have to be adapted to changes in clinical processes such as new tissue sampling methods to perform experiments.

9.3.1 "Tissue Banking"

For molecular research and the analysis of genomic DNA, mRNA, or proteins, high-quality human tissue is fundamental. Although for some of the morphology-based procedures, as well as for many applications of the PCR, FFPE tissue is adequate, most other diagnostic and research applications, which are based on intact genomic DNA, mRNA or protein, require frozen tissue samples, making the establishment of a frozen tissue bank a valuable diagnostic and research asset. In a clinical environment, usually several samples from one cancer patient—such as primary tumour, normal tissue, precursor lesions—are collected and stored at different time points with differing quality (paraffin tissue, serum, plasma) by different departments at different locations. Thus, it is necessary to consider standardization of the methods used to retrieve, freeze, store and catalogue tissue specimens by implementing standard operating procedures (SOP). Particularly, with regard to the collection and cryo-preservation of tissue, it is especially important to establish a rapid logistic chain for achieving optimum tissue preservation. This must be performed as soon as possible after ex-

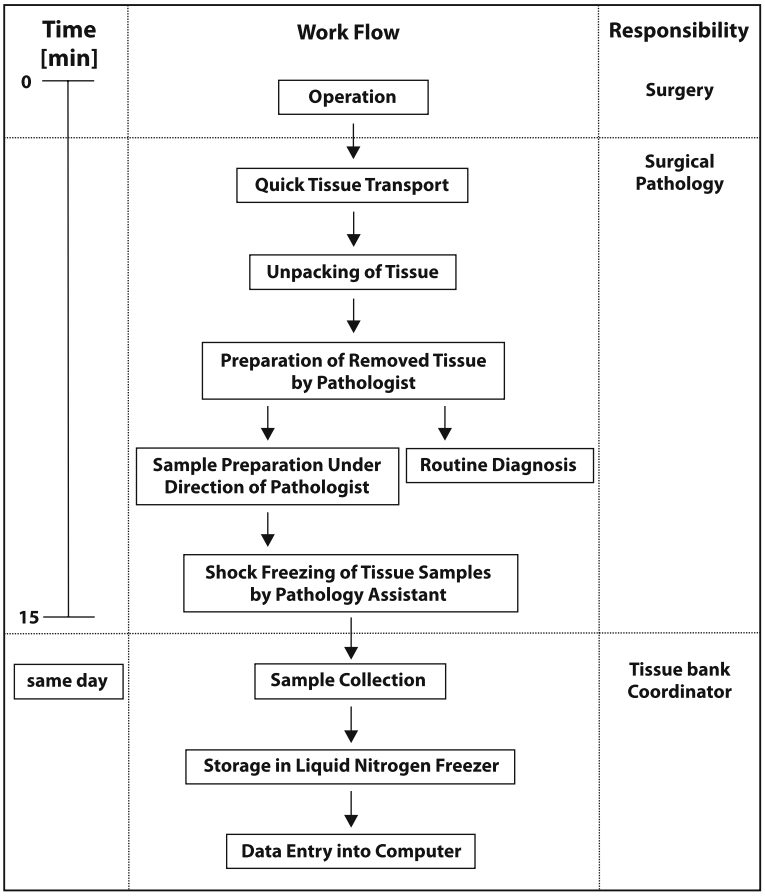


Fig. 9.5 Organizational structure of a frozen tissue bank. This diagram depicts the areas of responsibility for the surgeon, the surgical pathologist and the tissue bank co-ordinator and shows the workflow for sample collection and storage. Also depicted is the time after operation during which the tissue sample should be frozen to guarantee sufficient quality

cision of the tissue, optimally between 10 and 20 min afterwards. A responsible pathologist is required for the collection, maintenance of the tissue bank and its integration into the routine surgical pathology activities (Fig. 9.5).

Rapid freezing of the tissue samples can be achieved by several means and a variety of methods are in use in different laboratories. These include dropping freshly excised tissue into liquid nitrogen (−196 °C) or a cooled isopentane bath (−160 to −78 °C) after placing it into a 2.0-ml cryogenic vial without cryo-protection for subsequent use in procedures requiring the extraction of genomic DNA, mRNA or protein. To preserve tissue architecture and cytologic features for immunohistochemistry and in situ hybridization, freezing with a cryo-protectant such as Tissue-Tek OCT (VWR Scientific, Bridgeport, NJ) should be chosen. The mechanisms of tissue damage by freezing are complex; however, rapid cooling of the tissues retards the development and growth of ice crystals as well as abruptly halting enzymatic activity. Equally critical is the method used to store the frozen tissue for extended periods. Long-term storage of frozen tissues that may be used subsequently for molecular studies should be carried out at temperatures of −70 °C or lower. Long-term storage of the frozen tissue is recommended in the gas phase of liquid nitrogen which is under constant temperature surveillance. This is preferred because no chemical reactions take place below approximately −130 °C. Messenger RNA and proteins may undergo slow degradation in unfixed tissue stored at warmer temperatures, and growth of ice crystals is favoured, causing damage to the architectural and cytologic features of the tissue. Tissue sampled and stored under these conditions has been used successfully in molecular techniques analysing the transcriptome and proteome of tissue from breast cancer patients (Schütz et al. 2006; Neubauer et al. 2006). An alternative to cryo-preservation might be the storage of fresh tissue in high salt solutions (RNAlater, Ambion, Austin, TX) to preserve macromolecules. Ultimately the best way of monitoring the effectiveness of tissue storage conditions is to test the integrity of mRNA, because it is more labile than genomic DNA or proteins. Thus, the isolation of intact RNA is a good indication of how well the tissue has endured the frozen state (Fig. 9.6).

9.3.2 "Data Banking"

While tumour banking has been recognized for over a decade as a needed tool to advance the molecular science of oncogenesis and tumour progression (Naber et al. 1992; Naber 1996), only in the last 4–5 years has the linkage to clinical outcomes been regarded as crucial for achieving this goal (Qualman et al. 2004). Therefore, the development of databases represents an independent scientific field, which is separated from molecular and clinical research and demands high logistic and financial investment.

The introduction of databases into the clinic is increasingly important for the optimization of

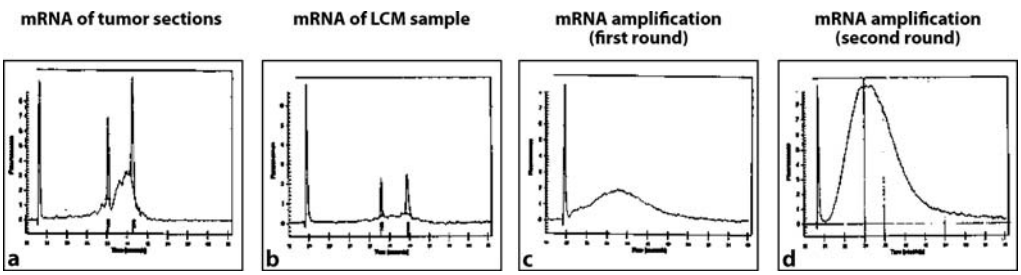


Fig. 9.6 a–d Quality control of tissue and analyte. Depicted is the control of rRNA integrity by capillary electrophoresis (Bioanalyzer, Agilent). **a** profile of rRNA from the whole tumour section; **b** profile of rRNA from laser dissected tumour cells; **c** profile amplified mRNA after first round of linear amplification; **d** profile amplified mRNA after a second round of linear amplification, before hybridization onto a microarray chip. The electropherograms in **a** and **b** show the 28S/18S rRNA peaks. The presence and a 2:1 ratio of the peaks indicate good quality. The *x-axis* indicates the time in seconds; *y-axis* indicates the measured fluorescence intensity

documentation, for quality assurance, to improve patient care, to gain reliable results and to control costs. Hence, database systems are not only important for bioinformatic data analysis but are required to develop a successful experimental design. In spite of complex database structures and functions, up to now no available database alone satisfies all medical requirements. In the field of specialized oncology databases are needed for the management of tissue banks for oncological research and care. To obtain exact results, small but defined sample collectives have to be selected out of tumour banks by combining several selection criteria. This attempt is impeded by the fact that different kinds of samples (e.g. fresh frozen, paraffin-embedded) and tissue pools, which are in general very heterogeneous because of different clinical progression and histology, might be needed. Therefore the combination of clinical data, histopathological data and data describing the course of the disease and therapy has to be enabled. In contrast to databases used in routine clinical practice for clinical documentation and quality assurance, tissue banks require tumour databases for molecular biological and scientific research. Therefore, not one single bank is needed but a collection of different banks, which are controlled by one database. The database has to provide the opportunity for flexible combined searches to gather the relevant information needed for stratification of samples. It must be pointed out that in addition to the logistic challenges collecting and storing tumour samples and associated data, there are numerous methodological, ethical and legal questions and challenges (Oosterhuis et al. 2003). In some countries co-operative or even nationwide tissue banks were established providing large numbers of annotated cancer specimens to investigators (Qualman et al. 2004; Schilsky et al. 2002; Melamed et al. 2004). Other nationwide tissue banks restrict the distribution of tissue to specifically funded projects. Recently, TumorAGENT, a new database, was developed for breast cancer research (Babel et al. 2006; Kurek et al. 2006). This dynamic relational database is unique in its structure and complexity; it has a flexible structure, architecture and user interface developed especially for the medical information and knowledge management field. It is able to collect,

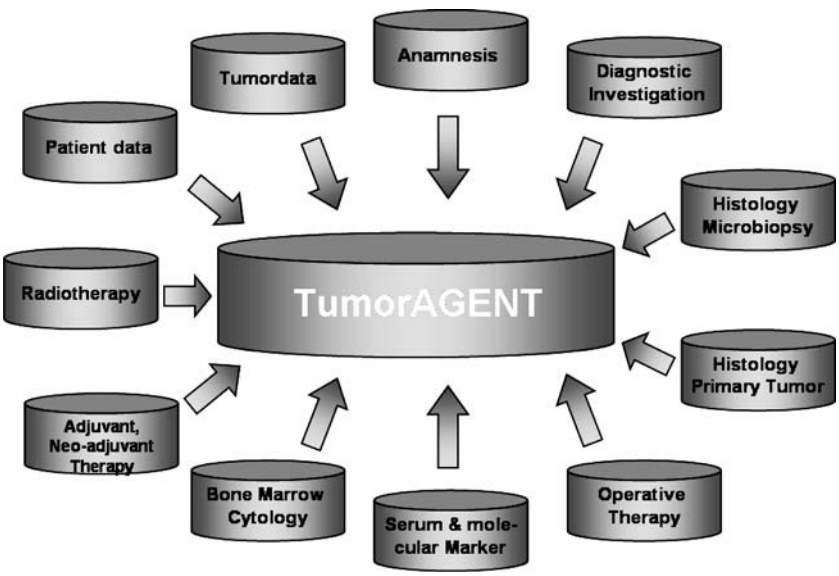


Fig. 9.7 Combined tissue and data management for oncologic research. Depicted are data modules implemented into TumorAGENT. This database consists of patient and sample management. In patient management all clinically relevant tumour data related to a patient can be collected and stored. Storage data of tissue samples are recorded in sample management. A flexibly designed search module permits data recall facilities over all data fields of patient and sample management. Even complex questions can be answered by the AND/OR relation

administrate, release and evaluate all relevant patient, tumour and sample data. The applied Web technologies offer maximum temporal and local availability. In TumorAGENT, approximately 800 attributes can be documented in about 100 forms and 30 entities, and it provides a flexible search module for molecular biological and scientific reporting (Fig. 9.7).

9.3.3 Sample Enrichment Strategies

The improvement of modern clinical screening programmes leads to the detection of tumours at an earlier stage. Consequently, profiling methods with increasing sensitivity have constantly to be improved. A challenging problem is the heterogeneity in tissue morphology accounting for the fact that the cell population of interest may constitute only a tiny fraction of the total tissue volume. For instance, breast cancer specimens are extremely heterogeneous with the tumour cells being mixed with many other cell types. Therefore, strategies for sample dissection, enrichment and amplification have to be employed to obtain specific expression profiles (Kunz and Chan 2004; Burgemeister 2005; Wulfkuhle et al. 2002). These techniques include e.g. magnetic beads, coupled with specific antibodies to label the cells of interest and to separate them by magnetic force. Alternatively, enrichment methods by the use of centrifugal forces applied to density gradients can be used (Jechlinger et al. 2003). For tissue sections, different macroscopic ("scratching") or microscopic techniques have been developed and should be applied by an experienced pathologist. One of these methods is microdissection assisted by a laser beam (Emmert-Buck et al. 1996; Wiltshire et al. 1995; Craven and Banks 2001; Bonner et al. 1997; Bichsel et al. 2000). It permits the isolation of single cells or single populations of cells from thin tissue sections (typically 5–10 µm in thickness) mounted on a glass slide (Fig. 9.8; Neubauer et al. 2006; Schütz et al. 2006).

It is reported that its application is able to reduce co-isolation of contaminating cells to as little as 0.6% (Nishidate et al. 2004). In one common type of this technology, laser capture microdissection (LCM), a narrow laser beam (7.5–30 µm in diameter) is fired at a heat-sensitive transparent polymer film on a cap that is in contact with the tissue. When the polymer is heated, it adheres to the cell(s) of interest and these cells are subsequently removed from the section when the cap is lifted (PixCell II LCM System, Arcturus, Mountain View, CA). An alternative approach is the laser microdissection and pressure catapulting (LMPC) system of PALM Microlaser Technologies (Bernried, Germany). In this system a pulsatile nitrogen laser is fitted to a modified research microscope and focussed on the sample. Specimens from different selected locations can first be laser microdissected and later catapulted by laser-induced pressure towards a collection device, such as a microtube cap (Poznanovic et al. 2005b).

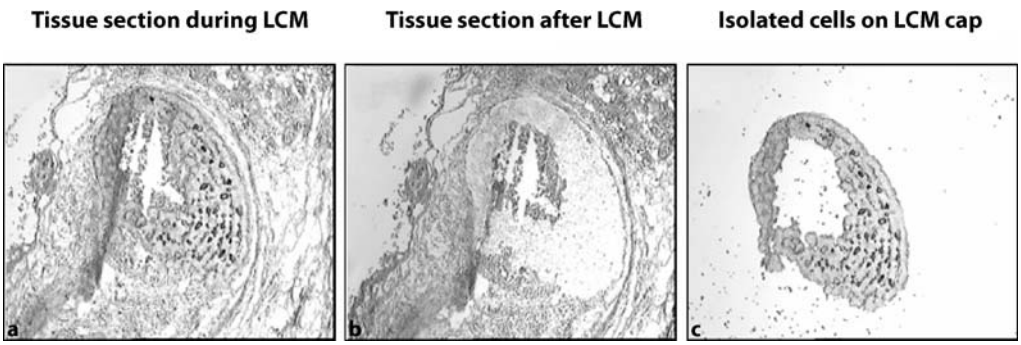


Fig. 9.8 a–c Tissue separation by Laser capture microdissection. Depicted is the isolation of epithelial cells of a breast ductal carcinoma in situ (DCIS). **a** The tissue section during LCM. Epithelial cells in the DCIS are dissected using a laser beam. The central necrotic area is not captured. **b** The tissue section after removal of the LCM cap with the epithelial being removed and the central necrotic area and the surrounding tissue being left on the slide. **c** The dissected epithelial cells on the LCM cap. The tissue section is stained with haematoxylin and eosin

The application of microdissection enables the analysis of cells from the same patient sample, e.g. different breast cancer progression stages, reducing the problem of genetic variability between individuals (Schütz et al. 2006). The other side of the coin is that the amounts of material that can be harvested by microdissection are exceptionally limited. This exacerbates the problems associated with sample yield and makes it difficult to standardize experiments, leading to reproducibility problems. Finally, experiments must be repeated to gain statistical sufficiency. The numerous replicates that are required to achieve sufficient sample sizes for acceptable statistical analysis is impossible for all but the most abundant proteins, which is exceedingly problematic with rare samples. For genomic and transcriptomic analyses, linear amplification or PCR-based amplification of the nucleic acid allows even single cells to be profiled (Schütz et al. 2006; Klein et al. 2002). But one has to acknowledge that the composition of the analyte might be biased, for instance, to the loss of the 5′-end of long transcripts. The issue of sample amount is ironically made worse by the success of contemporary clinical screening programmes, which detect tumours at an earlier stage. Therefore, high sensitivity detection methods are extremely desirable for the quantification of the small amounts of proteins often available in these samples, and the most successful studies represent the cutting edge of technology.

An important issue in using microdissection is to preserve the DNA, mRNA or the proteins. Therefore, special protocols must be established that allow quick staining of the tissue section using inhibitors for nucleases and proteinases and conserving its morphology. Eosin for example probably should be omitted or minimized from experiments investigating proteins (Craven et al. 2002). Maintaining the sections in a dehydrated state is crucial as well.

9.3.4 New Experimental Designs: Sample Pooling

Cells and their gene expression pattern are influenced by the environment, e.g. soluble factors, extracellular matrix proteins and the cell–cell communication they are deprived of when cultured. Ornstein et al. published a direct proteomic comparison between cultured clonally selected human prostate tumour cells and the patient-matched primary tumour epithelium from which they were derived (Ornstein et al. 2000). This revealed only a 25% similarity between the two protein populations even though they came from the same person. Therefore, molecular analysis of cells in their native tissue environment provides the most accurate picture of the in vivo disease state, with primary cell cultures established from fresh tissue being unable to duplicate the environment of cells in the actual tissue from which they are derived. Because of that, especially in experiments investigating clinical samples where the amount of tissue sample is particularly problematic, pooling of samples might be advantageous. Another reason for the pooling of samples is to reduce the costs of screening large numbers of samples. For proteomic analysis of microdissected clinical cancer samples, both of these situations apply, because proteins cannot be amplified. Therefore, the use of microdissection yields amounts of proteins that are difficult to reconcile with the need for greater amounts for 2D gels. In order to provide the amounts of protein obtainable from microdissection experiments on primary human samples from single patients with the amounts of protein required to achieve high quality 2D-PAGE results, a recent publication describes a sample pooling strategy (Neubauer et al. 2006). A disadvantage of pooling is that individual variations among pooled subjects are lost within a pool, so that all relevant clinical classifications are essential prior to pooling.

9.4 Conclusion

Due to technical progress of analytical methods and the knowledge of the human genome, the field of molecular medicine has become the basis for large-scale analysis of cancer cells on the genetic, transcriptomic and proteomic levels. These new technologies might transform the clinical practice of medicine, assisting with (1) the detection of cancer when it is at its earliest stage, even in the premalignant state, and (2) the individualization of treatments, which are two of the most important challenges of the post-genomic era. To accomplish these goals new biomarkers are needed to complement the existing histopatho-

logic markers being applied. This means that focussing efforts on one or a few platforms is unlikely to uncover all, or even the best, biomarkers of a disease. Therefore, different expression profiling platforms analysing genomic, transcriptomic and proteomic variations have to be used in a concerted manner. Regarding proteomic analysis, the methods available at the moment have sensitivities, a resolving power and a speed that already provide scientists with the possibility to analyse protein expression of complex systems such as breast cancer, with each approach having its strengths and weaknesses. Although these technologies are rapidly evolving they are still not as robust as those in the field of genomics, so that there are new strategies and technical improvements needed. On the other hand, even pre-analytical methods have to be improved and standardized to provide the optimal prerequisite for high-quality analysis and the comprehensive management of tissue, patient and experimental data. Further evaluations and the characterization of genomic, transcriptomic and proteomic variations may lead to the identification of biomarkers that can specifically be applied in clinical diagnoses, or might serve as drug targets.

References

Aebersold R, Goodlett DR (2001) Mass spectrometry in proteomics. Chem Rev 101:269–295

Alaiya A, Roblick U, Egevad L, Carlsson A, Franzen B, Volz D, Huwendiek S, Linder S, Auer G (2000) Polypeptide expression in prostate hyperplasia and prostate adenocarcinoma. Anal Cell Pathol 21:1–9

Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I (2003) A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 3:36–44

Alexander H, Stegner AL, Wagner-Mann C, Du Bois GC, Alexander S, Sauter ER (2004) Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. Clin Cancer Res 10:7500–7510

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM (2000) Distinct types of diffuse large B-cell lymphoma identified by gene-expression profiling. Nature 403:503–511

Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18:533–537

Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1:845–867

Arthur JM (2003) Proteomics. Curr Opin Nephrol Hypertens 12:423–430

Ayers M, Symmans WF, Stec J, Damokosh AI, Clark E, Hess K, Lecocke M, Metivier J, Booser D, Ibrahim N, Valero V, Royce M, Arun B, Whitman G, Ross J, Sneige N, Hortobagyi GN, Pusztai 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

Babel R, Fässler V, Geramani KN, Keckeisen M, Vartziotis D (2006) Interdisciplinary medical database network for the research of novel therapies and pharmaceuticals. International Special Topic Conference on Information Technology in Biomedicine, Ioannina

Baggerly KA, Morris JS, Coombes KR (2004) Reproducibility of the SELDI-TOF protein patterns in serum: comparing datasets from different experiments. Bioinformatics 20:777–785

Becker S, Cazares LH, Watson P, Lynch H, Semmes OJ, Drake RR, Laronga C (2004) Surfaced-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

Benohr P, Henkel V, Speer R, Vogel U, Sotlar K, Aydeniz B, Reiser A, Neubauer H, Tabiti K, Wallwiener D, Clare SE, Kurek R (2005) Her-2/neu expression in breast cancer—a comparison of different diagnostic methods. Anticancer Res 25:1895–1900

Bibikova M, Talantov D, Chudin E, Yeakley JM, Chen J, Doucet D, Wickham E, Atkins D, Barker D, Chee M, Wang Y, Fan JB (2004) Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. Am J Pathol 165:1799–1807

Bichsel VE, Petricion EF 3rd, Liotta LA (2000) The state of the art microdissection and its downstream applications. J Mol Med 78:B20

Bini L, Magi B, Marzocchi B, Arcuri F, Tripodi S, Cintorino M, Sanchez JC, Frutiger S, Hughes G, Pallini V, Hochstrasser DF, Tosi P (1997) Protein expression profiles in human breast ductal carcinoma and histologically normal tissue. Electrophoresis 18:2832–2841

Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69:89–95

Bonner RF, EmmertBuck M, Cole K, Pohida T, Chuaqui R, Goldstein S, Liotta LA (1997) Cell sampling: laser capture microdissection: molecular analysis of tissue. Science 278:1481–1483

Bons JA, Wodzig WK, van Dieijen-Visser MP (2005) Protein profiling as a diagnostic tool in clinical chemistry: a review. Clin Chem Lab Med 43:1281–1290

Bradshaw RA, Burlingame AL, Carr S, Aebersold R (2005) Protein identification: the good, the bad, and the ugly. Mol Cell Proteomics 4:1221–1222

Brentani RR, Carraro DM, Verjovski-Almeida S, Reis EM, Neves EJ, de Souza SJ, Carvalho AF, Brentani H, Reis LF (2005) Gene expression arrays in cancer research: methods and applications. Crit Rev Oncol Hematol 54:95–105

Bulyk ML, Gentalen E, Lockhart DJ, Church GM (1999) Quantifying DNA-protein interactions by double-stranded DNA arrays. Nat Biotechnol 17:573–577

Burgemeister R (2005) New aspects of laser microdissection in research and routine. J Histochem Cytochem 53:409–412

Butte A (2002) The use and analysis of microarray data. Nat Rev Drug Discov 1:951–960

Buyse M, Loi S, van 't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, Therasse P, Floore A, Amakrane M, Piette F, Rutgers E, Sotiriou C, Cardoso F, Piccart MJ; TRANSBIG Consortium (2006) Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 698:1183–1192

Cahill MA, Wozny W, Schwall G, Schroer K, Holzer K, Poznanovic S, Hunzinger C, Vogt JA, Stegmann W, Matthies H, Schrattenholz A (2003) Analysis of relative isotopologue abundances for quantitative profiling of complex protein mixtures labelled with the acrylamide/D3-acrylamide alkylation tag system. Rapid Commun Mass Spectrom 17:1283–1290

Cai Z, Chiu JF, He QY (2004) Application of proteomics in the study of tumor metastasis. Genomics Proteomics Bioinformatics 2:152–166

Carr S, Aebersold R, Baldwin M, et al (2004) The need for guidelines in publication of peptide and protein identification data: working Group on Publication Guidelines for Peptide and Protein Identification Data. Mol Cell Proteomics 3:531–533

Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Guiterrez MC, Elledge RM, Mohsin S, Osborne CK, Chamness GC, Allred DC, O'Connell P (2003) Gene expression profiling predicts therapeutic response to docetaxel [Taxotere(tm)] in breast cancer patients. Lancet 32:280–287

Chaurand P, DaGue BB, Pearsall RS, Threadgill DW, Caprioli RM (2001) Profiling proteins from azoxymethane-induced colon tumors at the molecular level by matrix-assisted laser desorption/ionization mass spectrometry. Proteomics 1:1320–1326

Chaurand P, Sanders ME, Jensen RA, Caprioli RM (2004) Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. Am J Pathol 165:1057–1068

Chen C, Boylan MT, Evans CA, Whetton AD, Wright EG (2005) Application of two-dimensional difference gel electrophoresis to studying bone marrow macrophages and their in vivo responses to ionizing radiation. J Proteome Res 4:1371–1380

Chen EI, Hewel J, Felding-Habermann B, Yates JR 3rd (2006) Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). Mol Cell Proteomics 5:53–56

Chen G, Gharib TG, Wang H, Huang CC, Kuick R, Thomas DG, Shedden KA, Misek DE, Taylor JM, Giordano TJ, Kardia SL, Iannettoni MD, Yee J, Hogg PJ, Orringer MB, Hanash SM, Beer DG (2003) Protein profiles associated with survival in lung adenocarcinoma. Proc Natl Acad Sci U S A 100:13537–13542

Choudhary J, Grant SG (2004) Proteomics in postgenomic neuroscience: the end of the beginning. Nat Neurosci 7:440–445

Cowherd SM, Espina VA, Petricoin EF 3rd, Liotta LA (2004) Proteomic analysis of human breast cancer tissue with laser-capture microdissection and reverse-phase protein microarrays. Clin Breast Cancer 5:385–392

Craven RA, Banks RE (2001) Laser capture microdissection and proteomics: possibilities and limitation. Proteomics 1:1200–1204

Craven RA, Totty N, Harnden P, Selby PJ, Banks RE (2002) Laser capture microdissection and two-dimensional polyacrylamide gel electrophoresis: evaluation of tissue preparation and sample limitations. Am J Pathol 160:815–822

Diamandis EP (2004) Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. J Natl Cancer Inst 96:353–356

Ding L, Chu S, Kunitake S, MacNair J, Cheung E, Ghosh M, Salunga R, Tuggle T, Wang W, Erlander MG, Sgroi D, Raja R (2004) Gene expression profiling of microdissected human breast cancer cells from formalin-fixed paraffin-embedded sections using microarrays. Proceedings of the American Association for Cancer Research. 95th Meeting, vol. 45, Cadmus Professional Communications, Linthicum, MD

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312:212–217

Eifel P, Axelson JA, Costa J, Crowley J, Curran WJ Jr, Deshler A, Fulton S, Hendricks CB, Kemeny M, Kornblith AB, Louis TA, Markman M, Mayer R, Roter D (2001) National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1–3, 2000. J Natl Cancer Inst 93:979–989

Ein-Dor L, Kela I, Getz G, Givol D, Domany E (2005) Outcome signature genes in breast cancer: is there a unique set? Bioinformatics 21:171–178

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, Liotta LA (1996) Laser capture microdissection. Science 274:998–1001

Espina V, Mehta AI, Winters ME, Calvert V, Wulfkuhle J, Petricoin EF 3rd, Liotta LA (2003) Protein microarrays: molecular profiling technologies for clinical specimens. Proteomics 3:2091–2100

Fan JB, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang B, Shen R, McBride C, Li HR, Fu XD, Oliphant A, Barker DL, Chee MS (2004) A versatile assay for high-throughput gene expression profiling on universal array matrices. Genome Res 14:878–885

Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7:85–97

Franzen B, Linder S, Alaiya AA, Eriksson E, Uruy K, Hirano T, Okuzawa K, Auer G (1996a) Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytokeratins. Br J Cancer 74:1632–1638

Franzen B, Auer G, Alaiya AA, Eriksson E, Uryu K, Hirano T, Okuzawa K, Kato H, Linder S (1996b) Assessment of homogeneity in polypeptide expression in breast carcinomas shows widely variable expression in highly malignant tumors. Int J Cancer 69:408–414

Gembitsky DS, Lawlor K, Jacovina A, Yaneva M, Tempst P (2004) A prototype antibody microarray platform to monitor changes in protein tyrosine phosphorylation. Mol Cell Proteomics 3:1102–1118

Gerling IC, Solomon SS, Bryer-Ash M (2003) Genomes, transcriptomes, proteomes: molecular medicine, its impact on medical practice. Arch Intern Med 163:190–198

Ginsburg GS, McCarthy JJ (2001) Personalized medicine: revolutionizing drug discovery and patient care. Trends Biotechnol 19:491–496

Gorg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665–3685

Goufman EI, Moshkovskii SA, Tikhonova OV, Lokhov PG, Zgoda VG, Serebryakova MV, Toropygin IY, Vlasova MA, Safarova MR, Makarov OV, Archakov AI (2006) Two-dimensional electrophoretic proteome study of serum thermostable fraction from patients with various tumor conditions. Biochemistry (Mosc) 71:354–360

Grizzle WE, Adam BL, Bigbee WL, Conrads TP, Carroll C, Feng Z, Izbicka E, Jendoubi M, Johnsey D, Kagan J, Leach RJ, McCarthy DB, Semmes OJ, Srivastava S, Srivastava S, Thompson IM, Thornquist MD, Verma M, Zhang Z, Zou Z (2003–2004) Serum protein expression profiling for cancer detection: validation of a SELDI-based approach for prostate cancer. Dis Markers 19:185–195

Gruber HJ, Hahn CD, Kada G, Riener CK, Harms GS, Ahrer W, Dax TG, Knaus HG (2000) Anomalous fluorescence enhancement of Cy3 and cy3.5 versus anomalous fluorescence loss of Cy5 and Cy7 upon covalent linking to IgG and noncovalent binding to avidin. Bioconjug Chem 11:696–704

Gygi SP, Han DK, Gingras AC, Sonenberg N, Aebersold R (1999a) Protein analysis by mass spectrometry and sequence database searching: tools for cancer research in the post-genomic era. Electrophoresis 20:310–319

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999b) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999

Gygi SP, Rochon Y, Franza BR, Aebersold R (1999c) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 18:1720–1730

Haab BB (2005) Antibody arrays in cancer research. Mol Cell Proteomics 4:377–383

Haab BB, Dunham MJ, Brown PO (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. Genome Biol 2:RESEARCH0004

Haas W, Faherty BK, Gerber SA, Elias JE, Beausoleil SA, Bakalarski CE, Li X, Villen J, Gygi SP (2006) Optimization and use of peptide mass measurement accuracy in shotgun proteomics. Mol Cell Proteomics 5:1326–1337

Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70

He QY, Chiu JF (2003) Proteomics in biomarker discovery and drug development. J Cell Biochem 89:868–886

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, 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

Hoheisel JD (2006) Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet 7:200–210

Hondermarck H, Vercoutter-Edouart AS, Revillion F, Lemoine J, el-Yazidi-Belkoura I, Nurcombe V, Peyrat JP (2001) Proteomics of breast cancer for marker discovery and signal pathway profiling. Proteomics 1:1216–1232

Hu Y, Zhang S, Yu J, Liu J, Zheng S (2005) SELDI-TOF-MS: the proteomics and bioinformatics approaches in the diagnosis of breast cancer. Breast 14:250–255

Hudelist G, Pacher-Zavisin M, Singer CF, Holper T, Kubista E, Schreiber M, Manavi M, Bilban M, 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

Hudelist G, Singer CF, Pischinger KI, Kaserer K, Manavi M, Kubista E, Czerwenka KF (2006) Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium. Proteomics 6:1989–2002

Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL (2004) A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet 36:299–303

Iwao-Koizumi K, Matoba R, Ueno N, Kim SJ, Ando A, Miyoshi Y, Maeda E, Noguchi S, Kato K (2005) Prediction of docetaxel response in human breast cancer by gene expression profiling. J Clin Oncol 23:422–431

Izzotti A, Bagnasco M, Cartiglia C, Longobardi M, De Flora S (2004) Proteomic analysis as related to transcriptome data in the lung of chromium(VI)-treated rats. Int J Oncol 24:1513–1522

Jansen MP, Foekens JA, van Staveren IL, Dirkzwager-Kiel MM, Ritstier K, Look MP, Meijer-van Gelder ME, Sieuwerts AM, Portengen H, Dorssers LC, Klijn JG, Berns EM (2005) Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. J Clin Oncol 23:732–740

Jarvinen AK, Hautaniemi S, Edgren H, Auvinen P, Saarela J, Kallioniemi OP, Monni O (2004) Are data from different gene expression microarray platforms comparable? Genomics 83:1164–1168

Jechlinger M, Grunert S, Tamir IH, Janda E, Ludemann S, Waerner T, Seither P, Weith A, Beug H, Kraut N (2003) Expression profiling of epithelial plasticity in tumor progression. Oncogene 22:7155–7169

Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, Thun MJ (2006) Cancer statistics. CA Cancer J Clin 56:106–130

Jenssen TK, Hovig E (2005) Gene-expression profiling in breast cancer. Lancet 365:634–635

Jessani N, Liu Y, Humphrey M, Cravatt BF (2002) Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. Proc Natl Acad Sci U S A 99:10335–10340

Jessani N, Niessen S, Wei BQ, Nicolau M, Humphrey M, Ji Y, Han W, Noh DY, Yates JR 3rd, Jeffrey SS, Cravatt BF (2005) A streamlined platform for high-content functional proteomics of primary human specimens. Nat Methods 2:691–697

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258:818–821

Klein CA, Seidl S, Petat-Dutter K, Offner S, Geigl JB, Schmidt-Kittler O, Wendler N, Passlick B, Huber RM, Schlimok G, Baeuerle PA, Riethmuller G (2002) Combined transcriptome and genome analysis of single micrometastatic cells. Nat Biotechnol 20:387–392

Koller A, Washburn MP, Lange BM, Andon NL, Deciu C, Haynes PA, Hays L, Schieltz D, Ulaszek R, Wei J, Wolters D, Yates JR 3rd (2002) Proteomic survey of metabolic pathways in rice. Proc Natl Acad Sci U S A 99:11969–11974

Komatsu S, Zang X, Tanaka N (2006) Comparison of two proteomics techniques used to identify proteins regulated by gibberellin in rice. J Proteome Res 5:270–276

Kondo T, Seike M, Mori Y, Fujii K, Yamada T, Hirohashi S (2003) Application of sensitive fluorescent dyes in linkage of laser microdissection and two-dimensional gel electrophoresis as a cancer proteomic study tool. Proteomics 3:1758–1766

Kunz GM Jr, Chan DW (2004) The use of laser capture microscopy in proteomics research—a review. Dis Markers 20:155–160

Kurek R, Babel R, Clare S, Endress A, Faessler V, Loeffler H, Neubauer H, Solomayer E, Wallwiener D, Fehm T (2006) TumorAGENT—an interdisciplinary tumor database for translational research in breast cancer. World Congress on Medical Physics and Biomedical Engineering 2006, Seoul

Laird PW (2005) Cancer epigenetics. Hum Mol Genet 14 Spec No 1:R65–76

Lander ES, Linton LM, Birren B, et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Li J, Zhang Z, Rosenzweig J, Wang YY, Chan DW (2002) Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin Chem 48:1296–1304

Li J, Steen H, Gygi SP (2003) Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. Mol Cell Proteomics 2:1198–1204

Lichter P, Tang CJ, Call K, Hermanson G, Evans GA, Housman D, Ward DC (1990) High resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. Science 247:64–69

Lilley KS, Friedman DB (2004) All about DIGE: quantification technology for differential-display 2D-gel proteomics. Expert Rev Proteomics 1:401–409

Liotta LA, Espina V, Mehta AI, Calvert V, Rosenblatt K, Geho D, Munson PJ, Young L, Wulfkuhle J, Petricoin EF 3rd (2003) Protein microarrays: meeting analytical challenges for clinical applications. Cancer Cell 3:317–325

MacBeath G (2002) Protein microarrays and proteomics. Nat Genet 32:526–532

MacBeath G, Schreiber SL (2000) Printing proteins as microarrays for high-throughput function determination. Science 289:1760–1763

Marshall E (2004) Getting the noise out of gene arrays. Science 306:630–631

Melamed J, Datta MW, Becich MJ, Orenstein JM, Dhir R, Silver S, Fidelia-Lambert M, Kadjacsy-Balla A, Macias V, Patel A, Walden PD, Bosland MC, Berman JJ (2004) The cooperative prostate cancer tissue resource: a specimen and data resource for cancer researchers. Clin Cancer Res 10:4614–4621

Mendrinos S, Nolen JD, Styblo T, Carlson G, Pohl J, Lewis M, 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

Meyer JM, Ginsburg GS (2002) The path to personalized medicine. Curr Opin Chem Biol 6:434–438

Miklos GL, Maleszka R (2004) Microarray reality checks in the context of a complex disease. Nat Biotechnol 22:615–621

Naber SP (1996) Continuing role of a frozen-tissue bank in molecular pathology. Diagn Mol Pathol 5:253–259

Naber SP, Smith LL, Wolfe HJ (1992) Role of the frozen tissue bank in molecular pathology. Diagn Mol Pathol 1:73–79

Neubauer H, Clare SE, Kurek R, Fehm T, Wallwiener D, Sotlar K, Nordheim A, Wozny W, Schwall GP, Poznanovic S, Sastri C, Hunzinger C, Stegmann W, Schrattenholz A, Cahill MA (2006) Breast cancer proteomics by laser capture microdissection, sample pooling, 54-cm IPG IEF, and differential iodine radioisotope detection. Electrophoresis 27:1840–1852

Nielsen UB, Cardone MH, Sinskey AJ, MacBeath G, Sorger PK (2003) Profiling receptor tyrosine kinase activation by using Ab microarrays. Proc Natl Acad Sci U S A 100:9330–9335

Nishidate T, Katagiri T, Lin ML, Mano Y, Miki Y, Kasumi F, Yoshimoto M, Tsunoda T, Hirata K, Nakamura Y (2004) Genome-wide gene-expression profiles of breast-cancer cells purified with laser microbeam microdissection: identification of genes associated with progression and metastasis. Int J Oncol 25:797–819

Nishizuka S, Charboneau L, Young L, Major S, Reinhold WC, Waltham M, Kouros-Mehr H, Bussey KJ, Lee JK, Espina V, Munson PJ, Petricoin E 3rd, Liotta LA, Weinstein JN (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

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

Ohno M, Miura T, Furuichi M, Tominaga Y, Tsuchimoto D, Sakumi K, Nakabeppu Y (2006) A genome-wide distribution of 8-oxoguanine correlates with the preferred regions for recombination and single nucleotide polymorphism in the human genome. Genome Res 16:567–575

Oosterhuis JW, Coebergh JW, van Veen EB (2003) Tumour banks: well-guarded treasures in the interest of patients. Nat Rev Cancer 3:73–77

Orchekowski R, Hamelinck D, Li L, Gliwa E, van-Brocklin M, Marrero JA, Vande Woude GF, Feng Z, Brand R, Haab BB (2005) Antibody microarray profiling reveals individual and combined serum proteins associated with pancreatic cancer. Cancer Res 65:11193–11202

Ornstein DK, Gillespie JW, Paweletz CP, Duray PH, Herring J, Vocke CD, Topalian SL, Bostwick DG, Linehan WM, Petricoin EF 3rd, Emmert-Buck MR (2000) Proteomic analysis of laser capture microdissected human prostate cancer and in vitro prostate cell lines. Electrophoresis 21:2235–2242

Ouyang Z, Takats Z, Blake TA, Gologan B, Guymon AJ, Wiseman JM, Oliver JC, Davisson VJ, Cooks RG (2003) Preparing protein microarrays by soft-landing of mass-selected ions. Science 301:1351–1354

Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351:2817–2826

Paik S, Kim CY, Song YK, Kim WS (2005) Technology insight: application of molecular techniques to formalin-fixed paraffin-embedded tissues from breast cancer. Nat Clin Pract Oncol 2:246–254

Patterson SD, Aebersold RH (2003) Proteomics: the first decade and beyond. Nat Genet 33 Suppl:311–323

Patton WF (2002) Detection technologies in proteome analysis. J Chromatogr B Analyt Technol Biomed Life Sci 771:3–31

Pawitan Y, Bjohle J, Amler L, Borg AL, Egyhazi S, Hall P, Han X, Holmberg L, Huang F, Klaar S, Liu ET, Miller L, Nordgren H, Ploner A, Sandelin K, Shaw PM, Smeds J, Skoog L, Wedren S, Bergh J (2005) Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res 7:R953–R964

Pawlik T, Hawke D, Liu Y, Krishnamurthy S, Fritsche H, Hunt K, 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

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. Nature 406:747–752

Piccart MJ, Loi S, Van 't Veer LJ, et al (2004) Multicenter external validation study of the Amsterdam 70-gene prognostic signature in node negative untreated breast cancer: are the results still outperforming the clinical-pathological criteria? San Antonio Breast Cancer Symposium, December 2004. Abstr 38

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20:207–211

Poland J, Cahill MA, Sinha P (2003) Isoelectric focusing in long immobilized pH gradient gels to improve protein separation in proteomic analysis. Electrophoresis 24:1271–1275

Poznanovic S, Schwall G, Zengerling H, Cahill MA (2005a) Isoelectric focusing in serial immobilized pH gradient gels to improve protein separation in proteomic analysis. Electrophoresis 26:3185–3190

Poznanovic S, Wozny W, Schwall GP, Sastri C, Hunzinger C, Stegmann W, Schrattenholz A, Buchner A, Gangnus R, Burgemeister R, Cahill MA (2005b) Differential radioactive proteomic analysis of microdissected renal cell carcinoma tissue by 54 cm isoelectric focusing in serial immobilized pH gradient gels. J Proteome Res 4:2117–2125

Pusztai L, Gregory BW, Baggerly KA, Peng B, Koomen J, Kuerer HM, Esteva FJ, Symmans WF, Wagner P, Hortobagyi GN, Laronga C, Semmes OJ, Wright GL Jr, Drake RR, 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

Qualman SJ, France M, Grizzle WE, LiVolsi VA, Moskaluk CA, Ramirez NC, Washington MK (2004) Establishing a tumour bank: banking, informatics and ethics. Br J Cancer 90:1115–1119

Ricolleau G, Charbonnel C, Lode L, Loussouarn D, Joalland MP, Bogumil R, Jourdain S, Minvielle S, Campone M, Deporte-Fety R, Campion L, 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

Rose JK, Bashir S, Giovannoni JJ, Jahn MM, Saravanan RS (2004) Tackling the plant proteome: practical approaches, hurdles and experimental tools. Plant J 39:715–733

Schena M, Heller RA, Theriault TP, Konrad K, Lachenmeier E, Davis RW (1998) Microarrays: biotechnology's discovery platform for functional genomics. Trends Biotechnol 16:301–306

Schilsky RL, Dressler LM, Bucci D, Monovich L, Jewell S, Suster S, Caligiuri MA, Kantoff PW, Compton C (2002) Cooperative group tissue banks as research resources: the cancer and leukemia group B tissue repositories. Clin Cancer Res 8:943–948

Schrattenholz A (2004) Proteomics: how to control highly dynamic patterns of millions of molecules and interpret changes correctly? Drug Discov Today 1:1–8

Schütz C, Bonin M, Clare SE, Nieselt K, Sotlar K, Fehm T, Solomayer E, Rieß O, Wallwiener D, Kurek R, Neubauer H (2006) Progression-specific Genes identified by gene Expression Profiling of matched Ductal Carcinomas in situ and Invasive Breast Tumors, combining Laser Capture Microdissection and Oligonucleotide Microarray Analysis. Cancer Res 66:5278–5286

Schwartz SA, Weil RJ, Johnson MD, Toms SA, Caprioli RM (2004) Protein profiling in brain tumors using mass spectrometry: feasibility of a new technique for the analysis of protein expression. Clin Cancer Res 10:981–987

Sgroi DC, Ma XJ, Ryan P, Smith B, Younger J, Baer TM, Haber D, Couch FJ, Goetz M, Ingle J, Suman V, Lingle W, Erlander MG (2004) Discovery of a gene expression signature that predicts clinical outcome in breast cancer patients treated with tamoxifen. Proceedings of the American Association for Cancer Research, 95th Meeting, vol. 45, Cadmus Professional Communications, Linthicum, MD

Shaw J, Rowlinson R, Nickson J, Stone T, Sweet A, Williams K, Tonge R (2003) Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. Proteomics 3:1181–1195

Sheehan KM, Calvert VS, Kay EW, Lu Y, Fishman D, Espina V, Aquino J, Speer R, Araujo R, Mills GB, Liotta LA, Petricoin EF 3rd, Wulfkuhle JD (2005) Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. Mol Cell Proteomics 4:346–355

Sieber SA, Mondala TS, Head SR, Cravatt BF (2004) Microarray platform for profiling enzyme activities in complex proteomes. J Am Chem Soc 126:15640–15641

Sitek B, Luttges J, Marcus K, Kloppel G, Schmiegel W, Meyer HE, Hahn SA, Stuhler K (2005) Application of fluorescence difference gel electrophoresis saturation labelling for the analysis of microdissected precursor lesions of pancreatic ductal adenocarcinoma. Proteomics 5:2665–2679

Sjogren S, Inganas M, Lindgren A, Holmberg L, Bergh J (1998) Prognostic and predictive value of c-erbB-2 overexpression in primary breast cancer, alone and in combination with other prognostic markers. J Clin Oncol 16:462–469

Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235:177–182

Snyderman R, Williams RS (2003) Prospective medicine: the next health care transformation. Acad Med 78:1079–1084

Somiari RI, Somiari S, Russell S, Shriver CD (2005) Proteomics of breast carcinoma. J Chromatogr B Analyt Technol Biomed Life Sci 815:215–225

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL (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 JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, 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

Speer R, Wulfkuhle JD, Liotta LA, Petricoin EF 3rd (2005) Abstract Reverse-phase protein microarrays for tissue-based analysis. Curr Opin Mol Ther 7:240–245

Sreekumar A, Nyati MK, Varambally S, Barrette TR, Ghosh D, Lawrence TS, Chinnaiyan AM (2001) Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. Cancer Res 61:7585–7593

Tlsty TD, Crawford YG, Holst CR, Fordyce CA, Zhang J, McDermott K, Kozakiewicz K, Gauthier ML (2004) Genetic and epigenetic changes in mammary epithelial cells may mimic early events in carcinogenesis. J Mammary Gland Biol Neoplasia 9:263–274

Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 1:377–396

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH (2002) Gene expression profiling predicts clinical outcome of breast cancer. Nature 31:530–536

Van de Vijver MJ, He YD, Van 't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R (2002) A gene expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999–2009

Velculescu VE, Vogelstein B, Kinzler KW (2000) Analysing uncharted transcriptomes with SAGE. Trends Genet 16:423–425

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M,

Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X (2001) The sequence of the human genome. Science 291:1304–1351

Vercoutter-Edouart AS, Lemoine J, Le Bourhis X, Louis H, Boilly B, Nurcombe V, Revillion F, Peyrat JP, Hondermarck H (2001) Proteomic analysis reveals that 14-3-3sigma is down-regulated in human breast cancer cells. Cancer Res 61:76–80

Vlahou A, Laronga C, Wilson L, Gregory B, Fournier K, McGaughey D, Perry RR, Wright GL Jr, Semmes OJ (2003) A novel approach toward development of a rapid blood test for breast cancer. Clin Breast Cancer 4:203–209

Von Eggeling F, Gawriljuk A, Fiedler W, Ernst G, Claussen U, Klose J, Romer I (2001) Fluorescent dual colour 2D-protein gel electrophoresis for rapid detection of differences in protein pattern with standard image analysis software. Int J Mol Med 8:373–377

Vuong GL, Weiss SM, Kammer W, Priemer M, Vingron M, Nordheim A, Cahill MA (2000) Improved sensitivity proteomics by postharvest alkylation and radioactive labelling of proteins. Electrophoresis 21:2594–2605

Wall DB, Kachman MT, Gong S, Hinderer R, Parus S, Misek DE, Hanash SM, Lubman DM (2000) Isoelectric focusing nonporous RP HPLC: a two-dimensional liquidphase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. Anal Chem 72:1099–1111

Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA (2005) Gene-expression profiles to predict distant metastasis of lymphnode-negative primary breast cancer. Lancet 365:671–679

Washburn MP (2003) Soft landing for protein chips. Nat Biotechnol 21:1156–1157

Washburn MP, Wolters D, Yates JR 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247

Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I (1995) Progress with gene-product mapping of the mollicutes: *Mycoplasma genitalium*. Electrophoresis 16:1090–1094

Whitelegge JP (2002) Plant proteomics: BLASTing out of a MudPIT. Proc Natl Acad Sci U S A 99:11564–11566

Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379:466–469

Wilson KE, Marouga R, Prime JE, Pashby DP, Orange PR, Crosier S, Keith AB, Lathe R, Mullins J, Estibeiro P, Bergling H, Hawkins E, Morris CM (2005) Comparative proteomic analysis using samples obtained with laser microdissection and saturation dye labelling. Proteomics 5:3851–3858

Wiltshire RN, Duray P, Bittner ML, Visakorpi T, Meltzer PS, Tuthill RJ, Liotta LA, Trent JM (1995) Direct visualization of the clonal progression of primary cutaneous melanoma: application of tissue microdissection and comparative genomic hybridization. Cancer Res 55:3954–3957

Wulfkuhle JD, Sgroi DC, Krutzsch H, McLean K, McGarvey K, Knowlton M, Chen S, Shu H, Sahin A, Kurek R, Wallwiener D, Merino MJ, Petricoin EF 3rd, Zhao Y, Steeg PS (2002) Proteomics of human breast ductal carcinoma in situ. Cancer Res 62:6740–6749

Wulfkuhle JD, Liotta LA, Petricoin EF (2003) Proteomic applications for the early detection of cancer. Nat Rev Cancer 3:267–275

Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, Roberts JR, Edgerton M, Gonzalez A, Nadaf S, Moore JH, Caprioli RM, Carbone DP (2003) Proteomic patterns of tumor subsets in non-small-cell lung cancer. Lancet 362:433–439

Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M (2000) Analysis of yeast protein kinases using protein chips. Nat Genet 26:283–289