The ongoing evolution of proteomics in malignancy

Amit S. Dhamoon, Elise C. Kohn and Nilofer S. Azad

Medical Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, 10 Center Dr. MSC 1500, Bethesda, MD 20892, United States

The complementary fields of genomics and proteomics offer insights into the molecular mechanisms of diseases. While genomics seeks to define our genetic substrate, proteomics explores the structure and function of proteins, which are the end effectors of our genes. Proteomics has been revolutionized in the past decade by the application of techniques such as protein arrays, two-dimensional gel electrophoresis, and mass spectrometry. These techniques have tremendous potential for biomarker development, target validation, diagnosis, prognosis, and optimization of treatment in medical care, especially in the field of clinical oncology. We will discuss innovations in proteomic technologies and highlight their prospective applications to patient care.

Introduction

Human disease often occurs as a result of disruptions in the DNA–mRNA–protein axis that underlies normal cellular physiology. Either inherited or spontaneous genetic mutations can lead to alterations in mRNA expression or protein function and trigger a disease state. An understanding of the molecular basis of disease often requires a multidisciplinary approach that elucidates the genetic anomaly as well as its functional consequences. The sequencing of the human genome in 2003 was a revolutionary step in the understanding of our genetic blueprint [1]. However, our phenotype is a manifestation of the proteome, the full complement of gene products, which execute the biological processes of the cell.

Proteomics is the study of protein structure and function. Unlike the relatively static nature of the genome, protein expression and function is dynamically regulated in health and dysregulation may result in disease. The genome contains 35 000 genes [1], but the proteome consists of greater than 1 million distinct protein species because of variations in splicing, processing, and post-translational modifications. Genetic polymorphisms [2] and high-throughput mRNA arrays [3] have been applied for prognostic and classification purposes of disease states. However, these methods cannot effectively identify disruptions in turnover rates, expression patterns, subcellular localization, quaternary structure, and post-translational modification states of proteins. Furthermore, each tissue of the body expresses a functionally distinct proteome that is uniquely regulated. High-throughput proteomic approaches have been used to develop biomarkers of disease, discover therapeutic targets, and provide diagnostic and prognostic information that may eventually guide treatment decisions, heralding a new era of personalized medicine. Proteomic technology is now used to elucidate the mechanistic bases of disease and predict response to treatment in clinical trials [4], with burgeoning potential for eventual clinical use [5]. Proteomic techniques and principles have been applied to widely disparate clinical disciplines. We will review recent advances in proteomics research and explore applications of these technologies primarily in the field of clinical oncology.

The proteomics toolbox

Proteomic technologies are varied and continuously evolving. Several low-throughput and high-throughput techniques have emerged on the forefront of proteomics research and are in various stages of clinical application. Traditional proteomic methods, such as immunohistochemistry (IHC), western blotting, immunoprecipitation, and ELISA, are regularly applied in the research arena and often used in the clinic for diagnostic purposes. The basic principles of the aforementioned techniques have been applied to more recent high-throughput proteomics methods. Tissue microarrays are created by formalin-fixation and paraffin embedding (FFPE) of slices of tissue cores onto a slide, which are then investigated by IHC techniques [6,7]. Minimal antibody is

Corresponding author: Azad, N.S. (azadn@mail.nih.gov)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2007.07.015

required and hundreds of samples on the same slide can be compared to each other, because of uniformity of antibody incubation conditions and washes. Separation techniques, including laser capture microdissection (LCM) [8,9] can be used to enrich for tissues of interest before analysis.

Other powerful high-throughput techniques for proteomic analysis include forward-phase and reverse-phase protein microarrays, depicted in Figure 1. Forward-phase arrays consist of antibodies that are immobilized on a slide in order to detect multiple proteins from a single patient. For example, antibodies that are specific for molecular players in key signaling cascades can be used to assay a tumor lysate. Conversely, with reverse-phase arrays, small amounts of many clinical samples are immobilized onto a slide and are then probed by an antibody specific for a protein of interest. Tight quality control and antibody validation and optimization are required for successful application of microarrays. Clinical implementation of these techniques could involve samples from already existing tissue banks. For example, Becker *et al.* have developed a protein extraction method to analyze proteins in FFPE tissues, a potentially rich source of information that was previously unavailable for molecular characterization [7].

The above techniques all require a target-driven approach, focusing on particular proteins or molecular pathways of interest. By contrast, two techniques, two-dimensional gel electrophoresis (2DGE) and mass spectrometry (MS), have revolutionized proteomic research. These techniques do not require pre-conceived assumptions about the identity or number of proteins characteristic for a particular disease. 2DGE, first developed in 1975, separates proteins initially according to charge and then by molecular weight in the second dimension [10]. This low-throughput technique is limited by the requirement of large amounts of starting material and poor sensitivity for low-abundance and low-molecular weight proteins and peptides [11]. 2DGE can be used to identify potential biomarkers, drug targets, or crucial mediators of

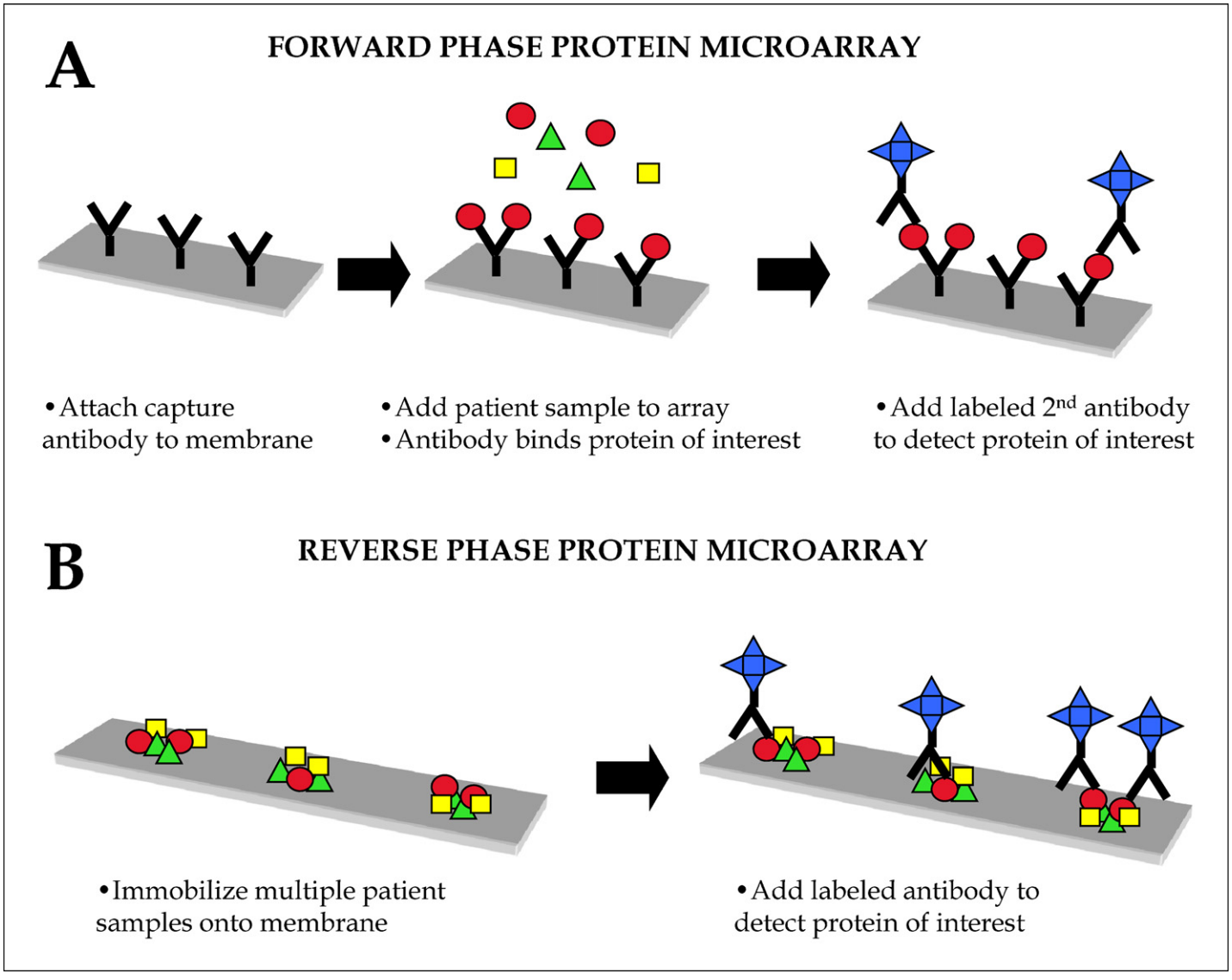

FIGURE 1

Schematic of forward phase and reverse phase tissue microarrays. (A) Forward phase protein microarrays utilize an antibody species that is immobilized onto a membrane or slide that can detect a specific protein in the tissue lysate. (B) Reverse phase protein microarrays consist of multiple tissue lysates that are immobilized onto a surface and then probed with an antibody in order to detect a specific protein of interest.

disease by comparing spot intensities between diseased and normal states. Peptides and proteins of interest can then be excised from the gel and identified. Further refinements in 2DGE technology, such as immobilized pH gradients, have allowed standardization of this technique across laboratories [12]. Also, in order to quantify differences in protein expression, samples can be pre-labeled with a unique fluorophore, mixed with other samples, and then run on the same gel [13,14]. This allows relative quantification of protein levels from different samples and avoids variability across multiple gels. 2DGE has served as a successful pre-clinical research tool to discover biomarkers of disease but because of its slow, labor-intensive process, it may be impractical for everyday clinical use.

Perhaps the most promising tool for proteomics research is mass spectrometry (MS). In MS, proteins and peptides are ionized, travel through an electric or magnetic field, and the mass-to-charge (m/z) ratios of the sample components are plotted [15] (see Figure 2). This technique serves as an extremely accurate scale for determining the molecular weight of individual peptides in complex protein mixtures. Mass spectrometers vary widely with respect to the means of sample introduction, ionization source, and resolution. Nevertheless, all mass spectrometers generate mass spectra of ions, consisting of the m/z ratio plotted on the x-axis and ion intensity on the y-axis. MS can be applied as a discovery tool to find differences in expression of individual proteins or profiles of proteins between diseased and healthy states. Alternatively, MS can be utilized to investigate the identity of proteins and peptides by comparing mass spectra with databases of known proteins. MS has the capability of identifying thousands of unique peptides from dozens of samples in the matter of hours. Because of the exquisite mass sensitivity of MS, small post-translational modifications, such as changes in phosphorylation state or disulfide bonding, can be accurately measured [16].

Before introduction into the MS, samples frequently are fractionated in order to enrich for proteins of interest in the analyte. In liquid chromatography mass spectrometry (LC–MS), complex clinical samples are separated by liquid chromatography and then introduced to the MS using electrospray technology [17]. An

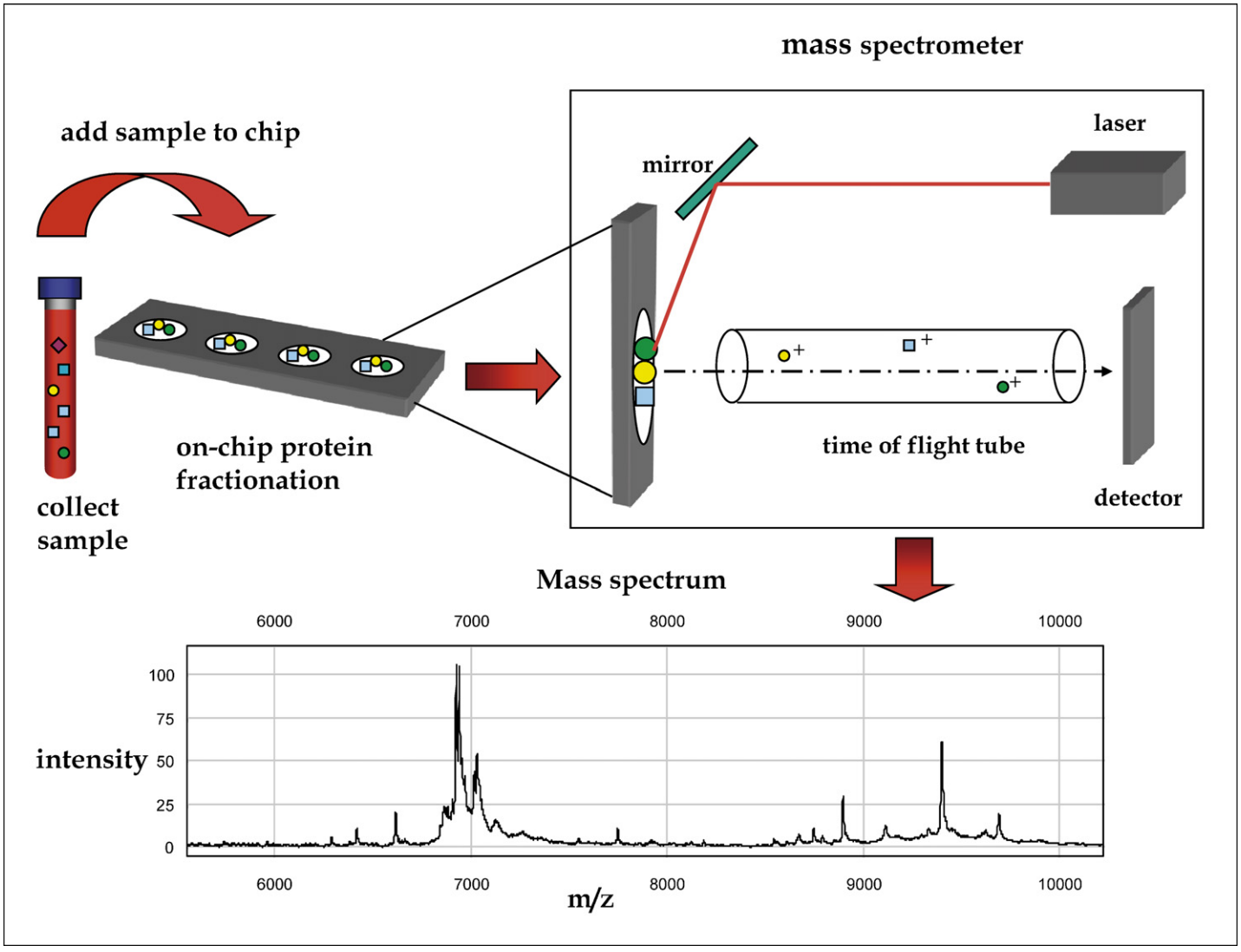

FIGURE 2

Schematic of SELDI-TOF–MS. Serum or plasma is collected from a patient and fractionated on a protein chip. Bound proteins are crystallized with organic acids and are then ionized by a laser. Ionized proteins and peptides travel through a time of flight tube and are separated by their mass-to-charge (m/z) ratio.

increasingly popular and easy-to-use MS platform is matrix-assisted laser desorption and ionization (MALDI), in which complex protein samples are spotted onto chips concurrently with an acidic organic matrix [18]. When a laser is fired at a sample spot on a plate, matrix molecules absorb the energy from the laser and prevent fragmentation of proteins and peptides during the ionization process. A variant of this technique, surface-enhanced laser desorption and ionization (SELDI), popularized by Ciphergen (Fremont, CA), allows for on-chip fractionation before interrogation by mass spectrometer [19] (see Figure 2). For example, if one is interested in studying negatively charged proteins in serum, such as albumin and its associated peptides, strong anion exchanger chips are used in order to selectively bind these proteins. Additional protein chips, including cation exchange, hydrophobic, and immobilized metal affinity capture can be implemented depending on the appropriate clinical application.

The SELDI-TOF–MS platform is conducive to clinical application because of the ease of sample preparation and potential for high-throughput automation; however, novel diagnostic biomarkers have yet to be discovered with this technique [20,21]. Most peaks of interest have been found to be isoforms and fragments of acute phase reactants, which may have limited specificity for a particular disease. Further, the mass spectrum is exquisitely sensitive to changes in pre-analytical and experimental variables, which may confound data analysis. SELDI–MS has intrinsically poor resolution in the high molecular weight range and is also not able to identify proteins that are less than 1 μg/ml in blood. Importantly, this technology does not have the resolution to identify potentially diagnostic low abundance low-molecular weight (LMW) peptides. It has been hypothesized that low abundance, LMW peptides in serum, a rich resource that has not been mined in depth, may contain diagnostic information [22]. In order to study this 'peptidome', high resolution mass spectrometers and novel fractionation techniques are required. Some laboratories have removed high-abundance proteins in order to study the remaining peptides for biomarker development [23]. Alternatively, other laboratories have hypothesized that high-abundance proteins serve as carrier molecules for small diagnostic peptides. In this approach, carrier proteins (including albumin and immunoglobulins) are concentrated and then LMW peptides are separated from the carrier molecule and analyzed [24]. Several alternative approaches, including the use of C18 magnetic beads and nanoporous surfaces may also have the potential to concentrate diagnostic carrier-bound peptides [25].

Mass spectrometry technology is continually evolving, with ongoing improvements in resolution, scanning speed, and ease of use [15,26]. Biomarker discovery may be improved by high-resolution instruments that have the ability to quantify differences in protein abundance from multiple samples using isotopic labeling [26]. These high-throughput proteomic technologies may eventually have vast potential for clinical use. The following sections will illustrate preliminary applications of these technologies for screening, diagnosis, and treatment of disease.

Applying proteomics for screening and diagnosis of disease

Considerable effort has been devoted to employ proteomic technologies in the development of disease biomarkers [27]. A biomarker is a quantifiable indicator that may be utilized to (1) detect early disease, (2) diagnose or rule out disease, (3) establish prognosis, or (4) follow treatment response. Unfortunately, development of single biomarkers has been hampered by insufficient sensitivity and specificity. For example, prostate-specific antigen (PSA) is routinely used in the clinic for early diagnosis of prostate cancer despite its poor sensitivity and specificity [28]. Other biomarkers, such as CEA, CA125, CA19-9, and others are most successfully applied for recurrence monitoring in patients with known disease. CA125 has been approved by the US Food and Drug Administration for use in monitoring response to treatment in women with epithelial ovarian cancer [29]. The poor specificity of biomarkers may lead to invasive, potentially harmful interventions while poor sensitivity can result in missed diagnoses.

It is increasingly recognized that identification of a single biomarker of disease may not be possible and that identifying panels of markers may be a more clinically useful paradigm. MS and 2DGE allow the screening of thousands of potential biomarkers at one time. The study of proteomic patterns, which is essentially a panel of dozens to hundreds of mass spectrometry peaks, has been applied to screen for disease states [5,30–32]. Proteomic signatures or patterns in breast [33], prostate [31], and ovarian cancers [30] among others, have been described using MS and 2DGE. Although diagnostic information has been obtained from proteomic patterns, critics may charge that differences between normal and disease states may be due to incidental differences caused by sample processing or, more importantly, overfitting of data instead of *bona fide* protein differences particular to the disease state [34,35]. For these reasons, scientists must validate their bioinformatic algorithms with independent, blinded patients sets and may attempt to determine the identity of diagnostic ions (Figure 3).

Biofluids, such as blood, urine, cerebrospinal fluid, nipple aspirate fluid, and others are rich sources of biomarkers of disease [33,34]. Clearly, a direct biopsy of an affected area is most likely to contain the needed diagnostic information but may also carry risk for patient complications. Biofluids in close proximity to the disease state may carry robust proteomic information with less risk to the patient.

Considerable attention has been focused on the analysis of the blood proteome, which may most accurately reflect the patient's overall physiological state, because the circulatory system is in constant contact with all tissues of the body. Acquisition of blood is minimally invasive and easily done, but study of blood proteins does pose several challenges. Plasma proteins span concentration gradients up to 10 orders of magnitude, while the dynamic range of current proteomic tools only spans 2–3 orders of magnitude [36]. Furthermore, about 20 proteins, including albumin, transferrins, and immunoglobulins, account for 99% of the protein content in blood [36]. It is postulated that biomarker information may be found in the remaining 1% of plasma protein content. Alternatively, others maintain that modified forms of abundant proteins may be excellent markers of disease. Furthermore, the plasma/serum proteome can rapidly change because of systemic effects that are unrelated to the disease of interest. Similar to initiatives to sequence the human genome, the Human Proteome Organization (HUPO) has embarked on cataloging components of human plasma [37–

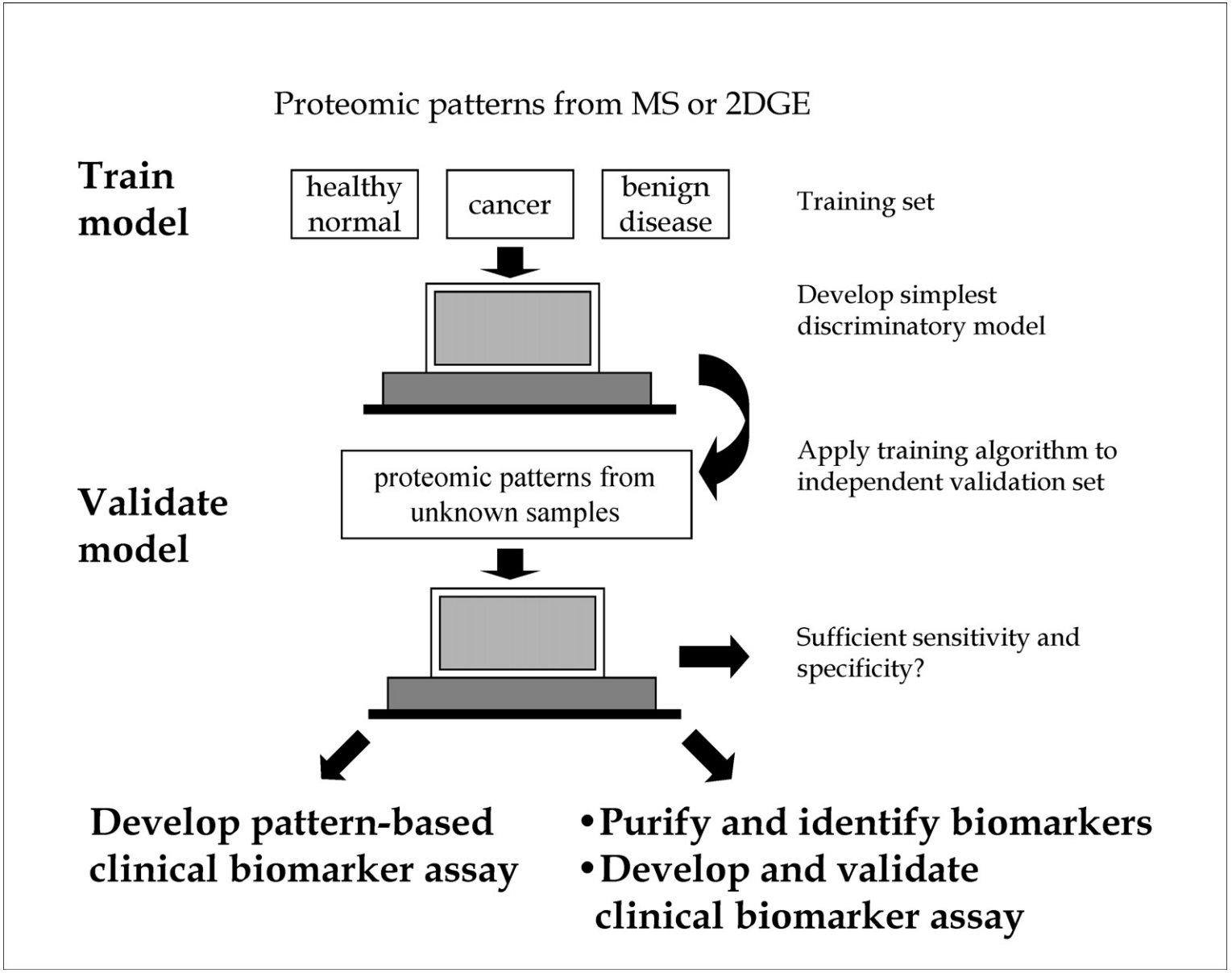

FIGURE 3

Schematic of workflow for MS and 2DGE biomarker discovery. After proteomic patterns are obtained by 2DGE or MS, bioinformatic algorithms are trained to distinguish between healthy, benign disease, and malignant states. These discriminatory algorithms are validated by testing on independent blinded sample sets. If the algorithm proves to have sufficient sensitivity and specificity, the biomarker pattern can be applied clinically, or alternatively, putative biomarkers can be identified. Once biomarkers are identified, traditional clinical assays can be developed.

39]. Secreted and transmembrane proteins from tissues have been found in the plasma proteome [40] and similarly, proteins from apoptotic and necrotic tissue would be expected to be found in the bloodstream as well. Nevertheless, finding tissue-specific proteins and peptides is especially difficult because of a tremendous dilution effect from the site of disease.

The first application of serum proteomics for MS based biomarker pattern development was by Petricoin *et al.* [30]. This groundbreaking and controversial paper used SELDI-TOF–MS to develop a bioinformatics algorithm that would differentiate serum collected from patients with ovarian cancer and from healthy unaffected women. This algorithm was then validated on an independent set of 116 patients with a sensitivity of 100% and a specificity of 95%. This approach has subsequently been used to develop putative diagnostics for other malignancies, including pancreatic [41], lung [42], liver [43], head and neck [6], and prostate cancer [31]. As described earlier, current PSA-based serum-screening tests for prostate cancer lack specificity and often result in invasive and often equivocal biopsies to rule out malignant processes. Several groups have used SELDI-TOF–MS and bioinformatics algorithms to develop serum proteomic profiles to distinguish prostate cancer from benign conditions, especially where the PSA is in the indeterminate range of 4–10 ng/ml [31,44]. Ornstein *et al.* refined this approach using a combination of high resolution mass spectrometry and a trained bioinformatics algorithm to discriminate prostate cancer from benign disease from an independent group of men (blinded to the investigators) with moderately elevated PSA levels [45]. Wang *et al.* [46] used a novel approach to develop a diagnostic test based on a protein microarray that screened for autoantibodies for prostate cancer [40]. The application of autoantibodies as a diagnostic tool for other cancers is ongoing [47].

Critics of pattern based serum proteomics charged that the identification of diagnostic MS peaks was required in order to assess their biological plausibility and to rule out artifactual confounders [34]. Using SELDI-TOF–MS, along with protein identification methods, several biomarkers, including haptoglobin [1], transferrin, transthyretin isoforms, apolipoprotein A1, inter-alpha trypsin inhibitor heavy chain 4 [40], and ferritin light chain [48] have been found to be associated with malignancies, though they have not been validated for widespread clinical utility or their specificity for a single cancer confirmed. Interestingly, diagnostic specificity may be obtained by studying the multiple isoforms or proteolytic fragments of a parent protein [40] that may be due to an underlying disease process. Although MS proteomic profiling may be a powerful tool in the screening of disease, more traditional assays should not be overlooked. Mor *et al.* developed a four-panel ELISA directed against four serum proteins, leptin, prolactin, osteopontin, and insulin-like growth factor-II that together had a sensitivity and specificity of 95% for detecting ovarian cancer [49]. This biomarker panel is being evaluated in a large prospective study. The limitation of the ELISA technique is that antibodies that are specific to multiple isoforms or proteolytic fragments of a protein are often not available. In this scenario, the exquisite sensitivity of antibodies and the mass sensitivity of MS can be combined in the technique, immunoaffinity-MS [50]. Antibodies are used to enrich for biomarkers of interest, which are then interrogated by MS in order to resolve the multiple isoforms of the protein of interest.

Considerable effort has focused on analysis of the blood proteome for disease screening and diagnosis. However, other biofluids that are proximal to the locus of disease have also been utilized for similar purposes. For example, urine collection is a safe non-invasive source of biomarker information for diseases of the kidney, bladder [51], and prostate [52,53]. Several groups have attempted to characterize the urine proteome by using 2DGE and MS identification of peptides [54,55]. Theodorescu *et al.* examined urine proteomic patterns utilizing capillary-electrophoresis-coupled MS for patients with urothelial cancer [56]. The authors initially developed a model to predict cancer based on 46 patients with urothelial carcinoma and 33 healthy volunteers. This algorithm was further refined using 366 urine samples from patients with urological malignancies, non-cancerous genitourinary (GU) disease, and healthy volunteers. Their diagnostic algorithm showed nearly 100% sensitivity and specificity for urothelial carcinoma. Furthermore, the authors discovered that fibrinopeptide A, a protein also associated with gastric and ovarian cancers, was one of the peptides in their diagnostic algorithm. Cerebrospinal fluid (CSF) is in direct contact with the central nervous system, and proteomic analysis has been applied to patients with neurodegenerative diseases [40], cancers of the CNS [57], and mental diseases [8,58]. Similarly, bronchoalveolar lavage fluid [59], synovial fluid [60], amniotic fluid [61], sputum [62], nipple aspirate fluid [63], and tears [64] have been interrogated for biomarker information.

In addition to screening for disease states, proteomic methods may also aid the clinician by providing crucial diagnostic and prognostic information. Pathologists render diagnoses after examining tissue samples using a combination of objective and subjective findings. Proteomic tools may provide more objective tools for tumor diagnosis as well as provide molecular characterization of the tumor that may guide treatment decisions.

In studying CNS tumors, Okamato *et al.* used a combination of tissue microdissection, 2DGE, and LC–MS to determine proteomic profiles of meningioma subtypes to aid in accurate grading of these tumors and subsequent identification of proteins of interest [65]. This study elegantly described the process of narrowing down the vague concept of a proteomic profile to specific markers of disease that can be applied to the clinic for diagnostics.

Oncologists have aggressively applied proteomic technologies to aid in cancer diagnosis and prognosis, but other fields are also benefiting from the burgeoning field of protein research. In the realm of infectious disease, tuberculosis is a highly prevalent but treatable disease that lies latent in one-third of the world's population. Death due to tuberculosis infection is avoidable if disease is diagnosed and treated early. Serum proteomic profiling using SELDI-TOF–MS was used to develop a proteomic fingerprint associated with TB infection, with a sensitivity and specificity of 93.5% and 95%, respectively [32]. Both the use of proteomic technologies of this study and the thoroughness of its design are worth noting. First, cases and controls were collected from multiple sites, minimizing the effects of institutional bias. Second, controls with a number of inflammatory conditions with symptoms that overlap with TB were also enrolled in order to develop a more discriminative diagnostic algorithm. After the diagnostic algorithm for serum was developed, it was then applied and validated on an independent blinded dataset.

Disease prognosis

Early knowledge of disease prognosis can guide informed decisions regarding treatment options for the patient. This is especially relevant in the field of oncology in which treatment for disease can be particularly onerous because of side effects. Goncalves *et al.* conducted SELDI-TOF–MS analysis of serum from women with breast cancer to develop a predictive algorithm for presence of metastasis [33]. Their model segregated patients with a good or poor prognosis and correctly predicted metastasis 83% of the time. Jacquemier *et al.* developed a model that predicted breast cancer prognosis using IHC with tissue microarrays from 552 patients [66]. In this target-driven approach, 21 proteins were able to separate patients with good versus poor prognosis in an independently validated set of 184 patients. Rakha *et al.* have studied a high-risk group of breast cancer patients that are hormone-receptor and HER2 negative in order to search for additional factors that would stratify patients with more aggressive tumors [67]. The authors examined samples from 1944 patients with invasive breast cancer who have been followed long term. Eleven proteins were assayed by tissue and several other factors, including age, pathological features, lymph node status, and histological grade were also studied. Basal cytokeratin expression found with tissue microarray analysis was an important prognostic factor in lymph node negative patients. Yanagisawa *et al.* examined tumor samples from patients with non-small cell lung cancer that were indistinguishable on the basis of clinical and pathological data [68]. The authors used tissue MALDI to develop proteomic profiles to provide prognostic information for these patients. Their bioinformatics algorithm successfully segregated patients with a mean survival time of 6 months from patients with a mean survival time of 33 months. In solid organ transplantation, Clarke *et al.* used SELDI–MS for proteomic profiling of urine to develop a non-invasive method for renal allograft prognostication with 83% sensitivity and 100% specificity [69]. Altogether, these techniques may allow modified dosing of therapy to maximize therapeutic potential and minimize side effects.

The application of proteomics for patient treatment

Understanding prognostic implications from proteomic technologies is an initial step in the development of personalized medicine, which is based on the particular molecular characterization of one's specific disease. Newer, target-driven chemotherapeutic agents have improved on older agents by modulating specific protein pathways instead of killing cells non-specifically. However, patients may respond to the same therapy in different ways. Proteomic techniques are being applied to fine-tune treatment and predict a patient's response to therapy or potential adverse events. At the NIH intramural research program, two phase II clinical trials were recently completed to evaluate the effect of targeted chemotherapy on patients with recurrent ovarian cancer. Gefitinib, a phosphorylation inhibitor of epidermal growth factor receptor (EGFR), and imatinib, an inhibitor of c-kit and platelet-derived growth factor receptor (PDGFR), were not found to have clinical benefit as single agents in this patient population. Needle biopsies were performed before and after treatment of imatinib and gefitinib, and reverse phase protein microarrays confirmed inhibition of drug targets and subsequent downstream effectors. A statistically significant correlation was found between clinical toxicity and side effects and the phosphorylation status of the drug targets in both trials [4,70]. Similarly, Xiao *et al.* used SELDI-TOF–MS to profile sera from patients with familial adenomatous polyposis to predict response to celecoxib therapy, which could allow patients to accurately weigh the costs and benefits of this cancer therapy [71]. Recently, Okano *et al.* identified biomarkers that would predict response to gefitinib treatment in patients with lung adenocarcinoma [72]. Gefitinib treatment is plagued by low-response rates and severe side effects to treatment. The 2DGE proteomic profile of adenocarcinoma tissue from 31 gefitinib responders was compared with 16 non-responders and a predictive algorithm based on expression levels of nine proteins was developed and validated on a small independent sample set. This study complements previous work that has highlighted the predictive potential of alterations in the EGFR gene [64,73] and mRNA analysis [74] in patients with lung cancer who are treated with gefitinib. The multidisciplinary approach offered by genotyping, functional genomics, and proteomics has provided insight into the molecular mechanisms of gefitinib sensitivity and serves as a model for integrating various disciplines for optimal patient care.

Conclusion

Proteomics offers exciting possibilities in the realm of clinical medicine. The field foretells advances in biomarker development, improved disease diagnosis and prognostic prediction, and drug development, with the ultimate promise of individualized personal therapy. Issues regarding reproducibility, reliability, standardization of experimental methodologies and analyses need to be optimized before final clinical application. Moreover, although the proteomics literature offers many examples of intriguing forays into novel areas of clinical application, few have taken the crucial step of validation of these techniques for widespread patient care. Multidisciplinary initiatives that utilize the strengths of distinct technologies, such as genomics and proteomics as well as the partnership of industry and academic research offer great potential for improvements in patient care in the near future.

Acknowledgements

This work was supported by the Intramural Research Program of the NCI, National Institutes of Health, Center for Cancer Research.

References

1 Collins, F.S. *et al.* (2003) A vision for the future of genomics research. *Nature* 422, 835–847

2 Zhao, X. *et al.* (2004) An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res.* 64, 3060–3071

3 Pollack, J.R. *et al.* (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* 23, 41–46

4 Posadas, E.M. *et al.* (2007) A phase II and pharmacodynamic study of gefitinib in patients with refractory or recurrent epithelial ovarian cancer. *Cancer* 109, 1323–1330

5 Azad, N.S. *et al.* (2006) Proteomics in clinical trials and practice: present uses and future promise. *Mol. Cell Proteomics* 5, 1819–1829

6 Becker, K.F. *et al.* (2006) Clinical proteomics: new trends for protein microarrays. *Curr. Med. Chem.* 13, 1831–1837

7 Becker, K.F. *et al.* (2007) Quantitative protein analysis from formalin-fixed tissues: implications for translational clinical research and nanoscale molecular diagnosis. *J. Pathol.* 211, 370–378
8 Emmert-Buck, M.R. *et al.* (1996) Laser capture microdissection. *Science* 274, 998–1001
9 Fuller, A.P. *et al.* (2003) Laser capture microdissection and advanced molecular analysis of human breast cancer. *J. Mammary Gland Biol. Neoplasia* 8, 335–345
10 O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021
11 Van den Bergh, G. and Arckens, L. (2005) Recent advances in 2D electrophoresis: an array of possibilities. *Expert Rev. Proteomics* 2, 243–252
12 Bjellqvist, B. *et al.* (1982) Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J. Biochem. Biophys. Methods* 6, 317–339
13 Von Eggeling, F. *et al.* (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
14 Lilley, K.S. and Friedman, D.B. (2004) All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev. Proteomics* 1, 401–409
15 Domon, B. and Aebersold, R. (2006) Mass spectrometry and protein analysis. *Science* 312, 212–217
16 Krueger, K.E. and Srivastava, S. (2006) Posttranslational protein modifications: current implications for cancer detection, prevention, and therapeutics. *Mol. Cell Proteomics* 5, 1799–1810
17 Qian, W.J. *et al.* (2006) Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol. Cell Proteomics* 5, 1727–1744
18 Hortin, G.L. (2006) The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. *Clin. Chem.* 52, 1223–1237
19 Seibert, V. *et al.* (2005) Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. *Brief Funct. Genomic Proteomic* 4, 16–26
20 Engwegen, J.Y. *et al.* (2006) Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. *Trends Pharmacol. Sci.* 27, 251–259
21 Poon, T.C. (2007) Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. *Expert Rev. Proteomics* 4, 51–65
22 Petricoin, E.F. *et al.* (2006) The blood peptidome: a higher dimension of information content for cancer biomarker discovery. *Nat. Rev. Cancer* 6, 961–967
23 Fountoulakis, M. *et al.* (2004) Depletion of the high-abundance plasma proteins. *Amino Acids* 27, 249–259
24 Zhou, M. *et al.* (2004) An investigation into the human serum ''interactome''. *Electrophoresis* 25, 1289–1298
25 Terracciano, R. *et al.* (2006) Selective binding and enrichment for low-molecular weight biomarker molecules in human plasma after exposure to nanoporous silica particles. *Proteomics* 6, 3243–3250
26 Ahn, N.G. *et al.* (2007) Achieving in-depth proteomics profiling by mass spectrometry. *ACS Chem. Biol.* 2, 39–52
27 Wagner, P.D. *et al.* (2004) Challenges for biomarkers in cancer detection. *Ann. N.Y. Acad. Sci.* 1022, 9–16
28 Bradford, T.J. *et al.* (2006) Molecular markers of prostate cancer. *Urol. Oncol.* 24, 538–551
29 Rustin, G.J. *et al.* (2006) Comparison of CA-125 and standard definitions of progression of ovarian cancer in the intergroup trial of cisplatin and paclitaxel versus cisplatin and cyclophosphamide. *J. Clin. Oncol.* 24, 45–51
30 Petricoin, E.F. *et al.* (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359, 572–577
31 Petricoin, E.F., 3rd *et al.* (2002) Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* 94, 1576–1578
32 Agranoff, D. *et al.* (2006) Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. *Lancet* 368, 1012–1021
33 Goncalves, A. *et al.* (2006) Postoperative serum proteomic profiles may predict metastatic relapse in high-risk primary breast cancer patients receiving adjuvant chemotherapy. *Oncogene* 25, 981–989
34 Diamandis, E.P. (2006) Serum proteomic profiling by matrix-assisted laser desorption–ionization time-of-flight mass spectrometry for cancer diagnosis: next steps. *Cancer Res.* 66, 5540–5541
35 Diamandis, E.P. (2004) Proteomic patterns to identify ovarian cancer: 3 years on. *Expert Rev. Mol. Diagn.* 4, 575–577
36 Anderson, N.L. and Anderson, N.G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell Proteomics* 1, 845–867
37 Omenn, G.S. (2005) Exploring the human plasma proteome. *Proteomics* 5, 3223–3225
38 Omenn, G.S. *et al.* (2005) Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 5, 3226–3245
39 States, D.J. *et al.* (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat. Biotechnol.* 24, 333–338
40 Fung, E.T. *et al.* (2005) Classification of cancer types by measuring variants of host response proteins using SELDI serum assays. *Int. J. Cancer* 115, 783–789
41 Koopmann, J. *et al.* (2004) Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. *Clin. Cancer Res.* 10, 860–868
42 Yang, S.Y. *et al.* (2005) Application of serum SELDI proteomic patterns in diagnosis of lung cancer. *BMC Cancer* 5, 83
43 Paradis, V. *et al.* (2005) Identification of a new marker of hepatocellular carcinoma by serum protein profiling of patients with chronic liver diseases. *Hepatology* 41, 40–47
44 Adam, B.L. *et al.* (2002) Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.* 62, 3609–3614
45 Ornstein, D.K. *et al.* (2004) Serum proteomic profiling can discriminate prostate cancer from benign prostates in men with total prostate specific antigen levels between 2.5 and 15.0 ng/ml. *J. Urol.* 172 (4 Pt 1), 1302–1305
46 Wang, X. *et al.* (2005) Autoantibody signatures in prostate cancer. *N Engl J Med* 353, 1224–1235
47 Cho-Chung, Y.S. (2006) Autoantibody biomarkers in the detection of cancer. *Biochim. Biophys. Acta* 1762, 587–591
48 Ricolleau, G. *et al.* (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
49 Mor, G. *et al.* (2005) Serum protein markers for early detection of ovarian cancer. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7677–7682
50 Nedelkov, D. (2006) Mass spectrometry-based immunoassays for the next phase of clinical applications. *Expert Rev. Proteomics* 3, 631–640
51 Pisitkun, T. *et al.* (2006) Discovery of urinary biomarkers. *Mol. Cell Proteomics* 5, 1760–1771
52 Downes, M.R. *et al.* (2006) Application of proteomic strategies to the identification of urinary biomarkers for prostate cancer: a review. *Biomarkers* 11, 406–416
53 Downes, M.R. *et al.* (2006) Urinary markers for prostate cancer. *BJU Int.*
54 Zerefos, P.G. *et al.* (2006) Characterization of the human urine proteome by preparative electrophoresis in combination with 2-DE. *Proteomics* 6, 4346–4355
55 Pieper, R. *et al.* (2004) Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics* 4, 1159–1174
56 Theodorescu, D. *et al.* (2006) Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. *Lancet Oncol.* 7, 230–240
57 Khwaja, F.W. *et al.* (2006) Proteomic analysis of cerebrospinal fluid discriminates malignant and nonmalignant disease of the central nervous system and identifies specific protein markers. *Proteomics* 6, 6277–6287
58 Romeo, M.J. *et al.* (2005) CSF proteome: a protein repository for potential biomarker identification. *Expert Rev. Proteomics* 2, 57–70
59 Magi, B. *et al.* (2006) Proteome analysis of bronchoalveolar lavage in lung diseases. *Proteomics* 6, 6354–6369
60 Liao, H. *et al.* (2004) Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. *Arthritis Rheum.* 50, 3792–3803
61 Ruetschi, U. *et al.* (2005) Proteomic analysis using protein chips to detect biomarkers in cervical and amniotic fluid in women with intra-amniotic inflammation. *J. Proteome Res.* 4, 2236–2242
62 Nicholas, B. *et al.* (2006) Shotgun proteomic analysis of human-induced sputum. *Proteomics* 6, 4390–4401
63 Pawlik, T.M. *et al.* (2005) Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry. *Breast Cancer Res. Treat.* 89, 149–157
64 Cappuzzo, F. *et al.* (2005) Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J. Natl. Cancer Inst.* 97, 643–655
65 Okamoto, H. *et al.* (2006) Comparative proteomic profiles of meningioma subtypes. *Cancer Res.* 66, 10199–10204
66 Jacquemier, J. *et al.* (2005) Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. *Cancer Res.* 65, 767–779
67 Rakha, E.A. *et al.* (2007) Prognostic markers in triple-negative breast cancer. *Cancer* 109, 25–32
68 Yanagisawa, K. *et al.* (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 362, 433–439

69 Clarke, W. *et al.* (2003) Characterization of renal allograft rejection by urinary proteomic analysis. *Ann. Surg.* 237, 660–664 discussion 664–5
70 Posadas, E.M. *et al.* (2004) A phase II clinical trial with proteomic profiling of imatinib mesylate in patients with refractory or relapsed epithelial ovarian cancer (EOC). In *Proceedings of the American Society of Clinical Oncology 40th Annual Meeting*, American Society of Clinical Oncology, pp. Abstr. 9651, June 5–8
71 Xiao, Z. *et al.* (2004) Serum proteomic profiles suggest celecoxib-modulated targets and response predictors. *Cancer Res.* 64, 2904–2909
72 Okano, T. *et al.* (2007) Proteomic signature corresponding to the response to gefitinib (Iressa, ZD1839), an epidermal growth factor receptor tyrosine kinase inhibitor in lung adenocarcinoma. *Clin. Cancer Res.* 13, 799–805
73 Hirsch, F.R. *et al.* (2003) Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *J. Clin. Oncol.* 21, 3798–3807
74 Kakiuchi, S. *et al.* (2004) Prediction of sensitivity of advanced non-small cell lung cancers to gefitinib (Iressa, ZD1839). *Hum. Mol. Genet.* 13, 3029–3043