

## REVIEW

# Strategies for plasma proteomic profiling of cancers

Gilbert S. Omenn

Departments of Internal Medicine, Human Genetics, and Public Health; Center for Computational Medicine and Biology, and Proteomics Alliance for Cancer Research, University of Michigan, Ann Arbor, MI, USA

Despite a voluminous literature on potential protein biomarkers and a compelling need for diagnostic tests based on biomarkers to detect cancers at much earlier, more treatable stages, progress has been limited. New methods and new instruments for analysis of differences in gene expression, gene methylation, and proteomics are being employed to try to accelerate the discovery phase. Given the heterogeneity of tumor mechanisms and the limitations of analytical methods, it is likely that a variety of strategies will be needed and will be complementary. That is the basis of this review of proteomic approaches. This article adopts a systems biology view, starting with mRNA transcripts in tumors and cultured tumor cells to detect mRNA overexpression, some of which will be correlated with protein overexpression. Some of those proteins may be secreted or released into proximal biofluids and plasma. Detection of low-abundance tumor proteins in the complex and dynamic mixture that is plasma requires combinations of increasingly powerful technologies. The biological amplification of protein signals through the immune system offers autoantibodies as potential biomarkers. Higher abundance proteins, including acute-phase reactants, may have practical value, especially if the proteins are modified as part of the cancer processes. Low molecular weight proteins, fragments, and peptides may offer complementary biomarkers. Promising biomarker candidates must be confirmed in independent studies. Then they must be submitted to higher-throughput methods practical for large-scale validation studies and, hopefully, for clinical and epidemiological applications. Standardized operating procedures for specimen handling, design and use of various reference standards, care to avoid bias and confounding, and guidelines for reporting findings and contributing datasets should enhance the prospects for predictive proteomic profiling of people at risk for cancers.

Received: May 3, 2006

Revised: July 20, 2006

Accepted: July 20, 2006

**Keywords:**

Biomarker / Cancer / mRNA transcripts Plasma / Systems biology

**Correspondence:** Dr. Gilbert S. Omenn, Departments of Internal Medicine, Human Genetics, and Public Health; Center for Computational Medicine and Biology, and Proteomics Alliance for Cancer Research, University of Michigan, Ann Arbor, MI, 48109-0656, USA

**E-mail:** gomenn@umich.edu

**Fax:** +1-734-647-8148

**Abbreviations:** CEA, carcinoembryonic antigen; CRP, cross-reactive protein; EGF, epidermal growth factor; EMT, epithelial-mesenchymal transition; ER, estrogen receptor; HCG, human chronic gonadotrophin; HER2/neu, human epidermal growth factor receptor 2; HUPO, Human Proteome Organization; ICAT, isotope-coded affinity tag; iTRAQ, isobaric-tagged relative and absolute quantitation; PR, progesterone receptor; PSA, prostate-specific antigen; SILAC, stable isotope labelling by amino acids in cell culture; TGF, transforming growth factor; VEGF, vascular epithelium growth factor

## 1 Introduction

Molecular biology has emerged during the past six decades with dramatic step functions in our knowledge. In 1944, Avery *et al.* demonstrated that DNA, not proteins, carries the information for inherited traits. In 1953, the x-ray diffraction results of Franklin and Wilkins inspired Watson and Crick to propose the hydrogen-bonded, double-helix structure for DNA and its replication. In the 1960s, Nirenberg and others demonstrated that a triplet code of nucleotide bases (A, T, G, C) specified, via translational mechanisms, the amino acid sequences in proteins. In the 1970s, came recombinant DNA and later the polymerase chain reaction (PCR) amplification of DNA. In the 1980s, the Human Genome Program was formulated and initiated, to greatly improve DNA and protein sequencing and synthesizing technologies, create

methods for mapping DNA segments, and then conduct huge-scale genome sequencing. Nearly complete human genome results were published 15–16 February 2001 in special issues of Science and Nature.

We have ushered in a “new world” of biology, genetics, and biotechnology—employing genome expression microarrays, comparative genomics, proteomics, and bioinformatics to seek to understand a great variety of diseases and eventually to transform medical care into a predictive, personalized, and preventive future model. Analytical instruments and bioinformatics have made experiments previously infeasible quite conceivable and increasingly feasible.

The aims of proteomic analyses of tumors and of plasma or serum from cancer patients are outlined in Table 1. The rationale for proteomic profiling is based on the fact that proteins are much closer to the pathophysiological changes in disease than are mRNAs. Up-regulated or modified proteins may themselves be molecular targets for drugs. Advances in fractionation of complex tissue and plasma protein mixtures, MS and antibody-based assays, and curated databases of proteins increasingly address the challenges of complexity, dynamic range, and incompleteness and uncertainty of protein identifications. Repositories with extensive datasets on normal protein findings include combined resources in UniProt and the International Protein Index managed by the European Bioinformatics Institute ([www.ebi.ac.uk](http://www.ebi.ac.uk)) and those from the Human Proteome Organization (HUPO) Plasma Proteome Project [1], which are publicly accessible at [www.ebi.ac.uk/pride](http://www.ebi.ac.uk/pride), [www.bioinformatics.med.umich.edu/hupo/ppp](http://www.bioinformatics.med.umich.edu/hupo/ppp), and [www.peptideatlas.org](http://www.peptideatlas.org).

**Table 1.** Aims of proteomics analyses of specimen from cancer patients

1. Profile tumor specimens for diagnosis and stratification of patients
2. Profile tumor specimens for prognosis with particular therapies
3. Discover and validate circulating proteins as biomarkers for earlier diagnosis
4. Apply such biomarkers to predict/monitor response to treatment and recurrence

Examining all genes and large numbers of proteins simultaneously, instead of one gene or one protein at a time, facilitates a systems biology approach to investigation of pathways, modules, regulatory networks, and responses to all kinds of perturbations [2]. In addition to yeast, *C. elegans*, Halobacterium and sea urchin models, these applications are most evident in cancer research [3], reflecting the somatic genetic basis for cancers and the desire to identify and validate new targets for more effective chemotherapy.

At the clinical and population level, common cancers remain largely unsolved medical problems. As shown in Table 2, according to the American Cancer Society 550 000

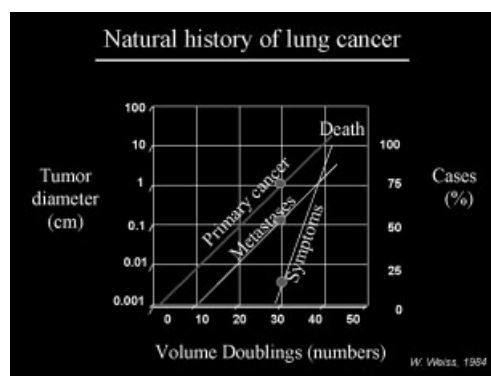
**Table 2.** Cancers remain largely unsolved medical problems

Cancer type	Incidence: new cases per year	Deaths per year, United States
Lung	171 900	157 200
Colorectal	147 500	57 100
Breast	211 300	39 800
Prostate	220 900	28 900
Pancreas	30 700	30 000
Ovarian	25 400	14 300
Total	2 000 000	550 000

Source: American Cancer Society

Americans die of various cancers each year. Lung cancers and pancreatic cancers remain extremely lethal, as deaths nearly equal incident cases, with short survival times. The poor outcomes of treatment for patients with lung cancers can be understood with Fig. 1, showing the relationship of tumor growth in doublings and volume to the late appearance of symptoms, later than the spread of metastases in most cases. The bottleneck on improving survival results with cancer therapies clearly lies in much earlier diagnosis. As shown in Table 3, even with present treatments, 5-year survival rates are far higher for patients whose lung, colorectal, breast, or prostate cancers are diagnosed while the tumor is still localized. Molecular biomarkers, primarily proteins, are the key to such earlier diagnoses, combined with imaging for localization.

An approach using molecular signatures and molecular mechanisms of cancers also offers a potential solution to the “pharmacogenomic nightmare” of the pharmaceutical companies. Currently, a drug that shows efficacy against, say, lung cancers, may be utilized in up to all of the 170 000 patients diagnosed annually with lung cancers in the U.S. alone. If molecular stratification of patients shows only 5–20% of patients have specific amplification or over-



**Figure 1.** Schema of growth of the primary lung cancer, developments of metastases, and very late appearance of symptoms sufficient to consider the diagnosis (kindly provided by Pierre Masson, Vanderbilt University).

**Table 3.** Detecting cancers early dramatically increases survival

Cancers	5-year survival rates	
	Late detection	Early detection
Lung	<10%	48%
Colorectal	8%	90%
Breast	16%	>95%
Prostate	33%	>95%

Source: American Cancer Society

expression or mutation of genes and receptors against which a particular drug is targeted, then the market would be reduced accordingly. This is true for human epidermal growth factor receptor 2 (HER2/neu) overexpression in certain breast cancers and mutated EGF-receptor tyrosine kinase domain in certain lung cancers. There is emerging evidence that similar mutations occur in some other tumors. If the same molecular signatures and mechanisms could be identified in, say, 5–20% of other common adenocarcinomas (colorectal, prostate, breast, lung, pancreas), then the combined market might be substantial and the efficacy:toxicity ratio would be greatly enhanced. There may still be other malignancies where the proportion of patients treatable with a particular agent will be very high, as with imatinib mesylate (Gleevec) targeted at bcr-abl translocations in Ph1-positive chronic myelogenous leukemias. However, the percentage of total leukemia patients is low. If the target is common in a tumor type, as appears to be the case in the majority of prostate cancers with the remarkable TMPRSS2/ETS fusion genes, which make ETS family transcription factors ERG and ETV1 androgen responsive [4], then new drugs, yet to be discovered, might have a benefit for that high proportion of patients. Conversely, if the target is common for multiple tumors, like VEGF-mediated angiogenesis, an anti-VEGF drug might be useful for patients with a variety of tumor types. Specific diagnostic tests are needed for each molecularly-targeted treatment.

This is an optimistic scenario. Molecularly targeted drugs are still few, despite vast investment in drug discovery and clinical trials. And new diagnostic tests have been very few, as well [5]. One reason is the marked heterogeneity of cancers even of the same organ of origin or similar histology. Another is the failure to characterize the positive predictive value for the test in relevant clinical scenarios before introducing the test for clinical use. Two of the most widely used tests—PSA and CA-125—have poor sensitivity and specificity, and thus low predictive value, for screening; however, they are useful for monitoring already-diagnosed patients for recurrence. According to Ludwig and Weinstein [6], despite 6000–8000 Medline entries for “cancer-related biomarkers” during each of the years 1999 through 2003, the total number of FDA-approved new tests was 2. Finding and validating new biomarkers has proved very difficult.

The cumulative roster of tests is limited to  $\alpha$ -fetoprotein for liver,  $\beta$ -HCG for testicular, CA-19–9 for pancreatic, CA-125 for ovarian, carcinoembryonic antigen (CEA) and EGFR for colon, KIT for GIST, thyroglobulin for thyroid, PSA for prostate; CA-15–3, CA-27–29, cytokeratins, ER/PR, and HER2/neu (aka ErbB2) for breast; and five for bladder cancers. Only Pap smears for cervical, PSA for prostate, and FISH/chromosome sites and NMP22 for bladder have been approved for screening, as opposed to monitoring already diagnosed patients for treatment response or recurrence. Each was approved as an “analyte-specific” assay [6].

Thus, there is a great need for new biomarkers, probably groups or panels of biomarkers to cover the expected heterogeneity of any common condition. One approach is to start from the biology of the cancers; another approach is to compare plasma or serum protein patterns in groups of patients and controls. With more sensitive instruments and emerging methods to fractionate and detect proteins and PTMs of proteins, these approaches should become more productive. No single approach is likely to detect the variety of paths that may yield useful biomarkers.

## 2 Discovering, confirming, and validating protein biomarkers

### 2.1 Starting from gene expression molecular profiles of tumors

There is a large literature on mRNA gene expression changes in tumors, compared with normal tissue and compared with non-malignant pathology in the same organs. A well-annotated resource for this literature is [www.oncomine.org](http://www.oncomine.org) [7] (Fig. 2). The changes in gene expression in tumors can be organized into phenomena, as in Hanahan and Weinberg’s “hallmarks of cancer” [8] (Table 4) and into functional groupings of differentially expressed mRNAs and proteins, as with prostate cancers (Fig. 3). These latter studies led to identification (and licensing) of hepsin,  $\alpha$ -methylacyl-CoA racemase (AMACR) [9], and the polycomb protein transcription factor EZH2 [10] as promising biomarker candidates for prostate cancer diagnosis and prognosis. EZH2 overexpression in the tumor by immunohistochemistry is highly correlated with metastatic *versus* localized prostate cancers and with poorer survival [10].

**Table 4.** Hallmarks of the pathogenesis of cancers [8]

1. Self sufficiency in growth signals
  2. Insensitivity to anti-growth signals
  3. Evasion of apoptosis
  4. Tissue invasion and metastasis
  5. Sustained angiogenesis
  6. Limitless replicative potential.
- Also, DNA repair and genomic instability

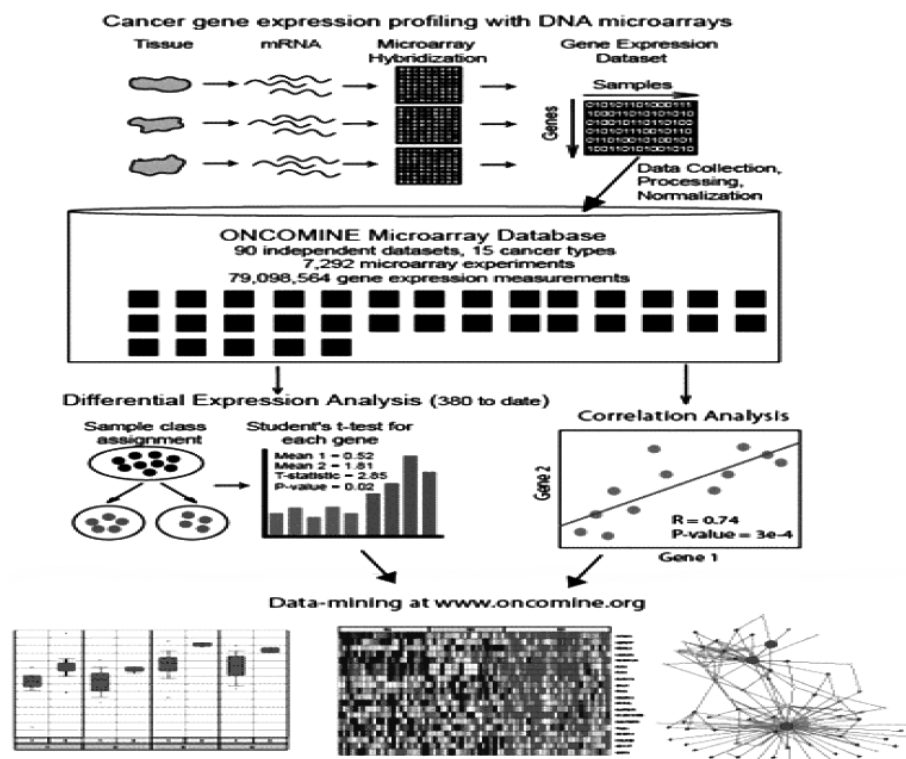


Figure 2. OncoPrint, a microarray database for cancer (www.oncoPrint.org).

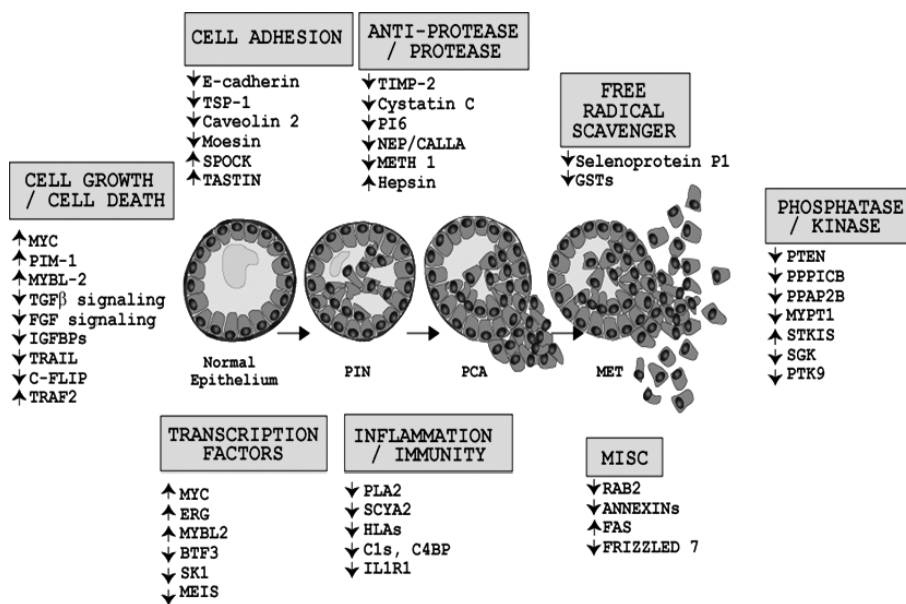


Figure 3. Summary of gene expression changes by protein functional classes during prostate cancer progression (provided by A. Chinnaiyan, University of Michigan).

These classifications of the molecular profiles of tumors move beyond lists of genes (or proteins) toward pathways and networks up-regulated or down-regulated over time during the initiation and progression of tumors. Such findings can be related to changes in cultured cells, xenografts, and patients with specific drug or siRNA interventions, and can be confirmed with direct detection and quantitation of

proteins in the tumor tissue. Variation in tumors due to inherited genomic differences, including single nucleotide polymorphisms; somatic chromosomal changes, including translocations; and epigenetic modifications, especially methylation (5-methyl-cytosine) of DNA, are additional important levels of characterization of tumors. However, while changes in some mRNA/protein pairs are pretty well

correlated, mRNA changes, overall, have surprisingly low correlation with changes in the corresponding proteins. This discrepancy presumably reflects the many post-translational events affecting protein structure, function, activation, and degradation. For example, Beer *et al.* [11] reported from a comparison of 76 human lung adenocarcinomas and 9 non-neoplastic lung specimens that 28 of 165 proteins identified by MS had a statistically significant correlation ( $r > 0.24$ ;  $p < 0.05$ ) with mRNA expression profiles in the same specimens; among all 165 proteins,  $r$  values ranged from  $-0.47$  to  $+0.44$ . Since proteins may have numerous isoforms, it is quite complex to compare mRNA and protein concentrations from very different methods. Different isoforms of presumably the same protein may differ in correlations with the mRNA. In the Beer study, 20 tumor-associated proteins were utilized to create a risk index for stage I patients, permitting elucidation of two subgroups with regard to survival, one with excellent 8-year survival, and the other with considerably poorer survival, more typical of more advanced lung cancer patients [12].

Varmabally *et al.* [13] used high-throughput immunoblotting of tissue extracts derived from prostate cancers; they identified 64 proteins altered in prostate cancer, compared with benign prostate hypertrophy, and 156 additional proteins altered in metastatic disease, compared with localized prostate cancer. Overall there was 48–64% concordance between protein and transcript expression level changes. Predictors of metastatic disease applied not just to prostate cancers, but to other solid tumors, as well.

It is useful to predict from the sequence features of corresponding proteins, and verify by direct measurement, which proteins will be secreted by tumor cells into the extracellular compartment and appear in lymph, other proximal fluids of specific organs, and, after huge dilution, the blood plasma. Targeted analysis of these biofluids for the expected proteins would depend then on the sensitivity of the method and the concentration of the protein, its stability and half-life, and any modifications of the protein that affect its detectability.

Direct analysis of tumors and of tumor cells fractionated by laser microdissection to permit analysis of the marked cellular heterogeneity of tumor masses may help reveal protein changes in tumors. There is a large literature of papers using 2-DE, now enhanced with paired analysis using fluorescent Cy dyes (DIGE). Most of the resulting papers report rather small numbers of proteins, as expected from the limited sensitivity of the gel-based method. However, Xiao *et al.* [14] identified 299 proteins in conditioned media of primary lung cancer or cell line cultures; 11 of 13 that were chosen for pilot ELISAs were detected in plasma.

After the tumor itself, it is logical to search for targeted proteins or conduct a general proteomic analysis of proximal biofluids [15]. These include urine for bladder and kidney cancers [16]; cerebrospinal fluid for brain tumors and other brain disorders [17]; nipple aspirate, ductal lavage, or interstitial fluid for breast tumors [18, 19]; pancreatic and bile

fluids for pancreatic and liver cancers; saliva for parotid and other salivary gland tumors; broncho-alveolar lavage or pleural effusion for lung cancers [20]; and prostatic aspirate, seminal fluid [21], tears, and other biofluids. However, we should note the complexity of transport or diffusion of molecules from the interstitium to the circulation. The endothelial cell is an effective barrier to most proteins above 40 kDa, so large proteins that are secreted may not actually make their way to the blood, at least not as intact proteins. Studies of such specimens, even in very small volumes, may be feasible with animal models of various human cancers. The Mouse Models of Human Cancers Consortium of the U.S. National Cancer Institute has multiple promising models, created with mutations discovered first in corresponding human cancers, such as the Apc (adenomatous polyposis coli) min/+ mouse with intestinal tumors (see [http://emice.nci.nih.gov/mouse\\_models](http://emice.nci.nih.gov/mouse_models)). There are Eastern and Western Consortia for Proteomics Technologies currently characterizing plasma samples from mouse models of colon, pancreatic, lung, and ovarian cancers.

## 2.2 Deducing potential protein biomarker candidates from *in vitro* studies of tumor cell lines

Despite the lack of cellular heterogeneity so important in tumors *in vivo*, it is useful to examine protein changes, especially up-regulation, in tumor cell lines *in vitro*. An example is the analysis of the epithelial-mesenchymal transition (EMT) [22] in TGF- $\beta$ -treated A549 human lung adenocarcinoma cells [23]. This system is a model for the transition of tumors from localized to an invasive and metastatic phenotype, as bioassays show striking increases in tumor cell motility and invasiveness. There are well-established protein biomarkers for the EMT, namely down-regulation of E-cadherin, up-regulation of N-cadherin, and up-regulation of vimentin. Keshamouni *et al.* [23] employed the iTRAQ method with isobaric tag reagents [24]. The initial studies compared protein profiles after 72 h of treatment with 5 ng/mL TGF- $\beta$  versus baseline, using a double-duplex design with the four tags, thereby facilitating significant improvement in the normalization of quantitative results. With the ABI 4700 instrument for LC-MS/MS, 29 proteins were identified as up-regulated and 22 as down-regulated (51 total vs. 16 expected by chance). Proteins involved in cytoskeletal reorganization, including actin, tubulin, and tropomyosins, and proteins involved in cell motility, including integrin  $\beta$ -1, filamins, cofilin, Hsp27 (HspB1), transglutaminase, and the ezrin-radixin-moesin complex were up-regulated, while the down-regulated proteins were more involved in cellular metabolism. The cadherins and vimentin were not detected, which may reflect concentration below the threshold for detection or interference with detection of their peptides. Vimentin (but not cadherins) and other additional proteins were identified in complementary proteomic analyses with DIGE and SILAC methods, as reported at the US HUPO March 2006

meeting. Elaborate combined time-course experiments with both Affy analysis of mRNA and iTRAQ analysis of proteins are pending. GeneGo Metacore pathways software tools were applied to the proteomics findings, showing significant connections among the up-regulated proteins [23]. siRNA knockdown of HspB1 remarkably dissected EMT into two phases, differentiating the down-regulation of E-cadherin from the up-regulation of N-cadherin and vimentin. Analysis of the literature revealed extensive evidence that many of the up-regulated proteins are involved in the biological processes modeled here (see Table 5). Analyses of culture media will be performed to determine which, if any, of these proteins are secreted or released in sufficient amounts to be candidates for detection after dilution in the plasma.

Many other interesting cell culture experiments have reported major changes in protein expression. Lin *et al.* [3] combined massively parallel signature sequencing (MPSS) of mRNA transcripts with isotope-coded affinity tag (ICAT) of the proteome to conduct a systems biology study of androgen-perturbed LNCaP cells. The combined patterns show functional differences between androgen-dependent and androgen-independent prostate cancer cells. Gu *et al.* [25] applied amino acid ICAT tagging to track p53-induced apoptosis in colorectal cancer cells. Protein expression was altered in multiple functional categories: cell cycle, membranes, reactive oxygen species, glycolysis, ATP transport, MMP induction, organelle cross-talk, and chaperones. Witt *et*

*al.* [26] utilized high throughput recombination-based cloning of 1000 cDNAs relevant to breast cancers and a translational system to screen for proteins that mediate phenotypic changes in three cell-based assays of mammary-10A tumor cells expressing an ErbB2 variant; the assays were EGF-independent proliferation, disruption of acinar morphogenesis, and cell migratory behavior independently or cooperatively with ErbB2 activation.

Nishizuka *et al.* [27] characterized the NCI-60 cancer cell lines with cDNA arrays, re-sequencing of the clones, and corroboration using Affymetrix oligonucleotide arrays. RP proteomic arrays evaluated the selectivities of candidate biomarkers at the protein level. Then tissue arrays were used to demonstrate similar selectivity in clinical tumor specimens, differentiating colon cancers from ovarian cancers and from tumors of unknown primary site [28].

Technological advances are moving toward much more sensitive detection of phosphorylation and other PTMs of proteins in very small numbers of cells, using “ultra-microarrays”, illustrated with PSA and IL-6 in <100 LNCaP cells [29]. Up to 17 properties, including signaling pathways, have been measured in single cells with flow cytometry [30]. The flow cytometry methods may turn heterogeneity of tumors into an advantage, as different types of cells can serve as comparisons or controls for the tumor cells or inflammatory cells detected. These methods might also be enhanced sufficiently to detect rare cancer cells in the circulation, as well as cancer stem cells within the tumors.

**Table 5.** Correlation of TGF-beta-induced up-regulated proteins in A549 human lung adenocarcinoma cells with roles reported in human cancers and metastasis-related processes (from [23])

Protein name	Human cancer	Lung cancer	Metastatic processes			
			Adhesion	Migration	Invasion	Metastasis
Tropomyosins	+	+		+		
Calgizzarin	+					
Filamin A, B, and C	+		+	+		+
Integrin beta 1	+	+	+	+	+	+
HSPB1	+	+		+		+
Non-muscle myosin heavy polypeptide-9			+	+		
Transglutaminase 2	+	+	+	+	+	+
Transgelin 2	+			+		
Myosin alkali light chain	+		+	+		+
Radixin	+	+	+	+		
Moesin	+	+	+	+		
Desmoyokin	+		+			
Cofilin	+		+	+		
Glutathione transferase	+					
14-3-3 zeta/delta	+	+		+		+
Keratin	+	+		+		
Annexin A2	+	+		+		+
Actin, cytoplasmic	+		+	+		
Tubulins	+	+		+		
Histone H2A	+					

### 2.3 Detecting tumor proteins in plasma or serum

There are two general approaches: searching for proteins already suspected or nominated as biomarker candidates from tumor and tumor cell studies (above), and comparing plasma or serum from cancer patients against specimens from very well-matched non-cancer patients and normals to detect proteome differences. Blood obtained by venepuncture is by far the most accessible human specimen, least invasive, and feasible to monitor over long periods of time. The blood plasma may capture proteins released from all organs and tissues in health and disease. However, plasma or serum presents certain daunting challenges: the enormous complexity in numbers of proteins and protein isoforms; the range of concentration from albumin at 40 mg/mL to cytokines at 1 ng/mL and tissue proteins at even lower concentrations; the dynamic changes in concentration, structure, and function as a result of physiological, pathological, and pharmacological triggers; and the necessity to tradeoff extensive fractionation of intact proteins and of peptides after tryptic digestion against higher throughput of less fractionated specimens [1].

Proteins in the circulation that were released by tumors, especially early-stage tumors, would be expected to have very low concentrations. Proteins are susceptible to cleavage in the circulation and during venepuncture, processing, and storage. They are likely to be excreted by the kidney after filtration through the glomeruli, if molecular weight is not greater than 40 kDa. Low molecular weight proteins or protein fragments may be retained, however, if bound to albumin or other highly abundant proteins.

#### 2.3.1 Targeting expected low-abundance tumor-associated proteins

Existing protein biomarkers like prostate-specific antigen (PSA), CA-125 for ovarian cancers, CA-19-9 for pancreatic cancers, CEA for gastrointestinal cancers, and C-reactive protein (CRP) for inflammatory conditions and acute cardiovascular events are detected and quantitated with specific immunoassays. Each assay required extensive work to optimize conditions and reduce coefficient of variation. In principle, the same kinds of assays can be applied to other tumor-associated proteins shown to be secreted or released by tumors and tumor cells. To study many proteins simultaneously, however, MS or antibody array methods are needed [31]. Simultaneous optimization for many assays has not yet been accomplished.

Anderson and Hunter [32] have introduced a quantitative electrospray LC-MS/MS multiple reaction monitoring (MRM) assay for pre-selected, protein-specific tryptic peptides. Their proof-of-principle publication detected a few dozen high and medium abundance proteins, from albumin (40 mg/mL) to fibronectin (300 µg/mL) and a peptide from L-selectin (670 ng/mL), quantitated with an ABI 4000 Q TRAP hybrid triple quadrupole/linear IT instrument. The

peptide from L-selectin (AEIEYLEK) appears twice in the HUPO PPP database, from two different labs, using two different citrate-anticoagulated reference specimens ([1]; see PPP website). Peptides vary enormously in their detectability in the mass spectrometer, so shrewd judgment and trial-and-error are needed to maximize sensitivity of this approach. Enrichment with stable isotope standards and capture with anti-peptide antibodies (SISCAPA) can extend the range of detection with this approach. These methods can be coupled with survey methods and with miniaturization using beads [33] or microarrays [34]. The instrumentation is widely available for assays of metabolites.

Aebersold has described a proteotypic peptide, isotope-labeling approach focused on glycoproteins. This method is well suited for spiking and quantitatively analyzing potentially very large numbers of proteins [35, 36]. Enrichment of glycoproteins in plasma has the special benefit of eliminating albumin, which has no sugars attached. The Hancock and Aebersold labs, using complementary lectin-binding and hydrazide methods, identified 264 glycoproteins with HUPO reference specimens [1, 37]. This proteotypic peptide/glycosite method has particularly high potential for identifying and quantifying proteins in a high throughput mode.

#### 2.3.2 Detecting low-abundance tumor proteins, protein fragments, or peptides bound to abundant proteins

Petricoin *et al.* [38] described a method of biomarker amplification and harvesting from carrier molecules, colorfully termed “molecular mops” or “molecular sponges” with their “diagnostic cargo”. The concept is that protein fragments or peptides, which would be eliminated through the kidney if not bound, are retained and thereby amplified above their concentration emerging from the tumor site, by being bound to albumin or other highly abundant plasma proteins. A variety of methods can be applied to detect the proteins, ranging from a look-up table of  $m/z$  ratios and TOF for previously sequenced low molecular weight protein fragments to MS/MS identification of previously unidentified fragments. Some of these fragments may be low-abundance disease-related isoforms of high abundance proteins. The new Orbitrap mass spectrometer seems to be well suited to these analyses. Tryptic digestion can be bypassed, given the mostly low  $M_r$  of fragments or peptides. However, the risk of errors in protein identification is increased when proteolysis is due to a diverse set of endogenous enzymes, rather than limited to tryptic peptides.

#### 2.3.3 Differences in the low molecular weight serum peptidome

Villanueva *et al.* [39] published a surprising approach based on what they interpreted to be primarily *ex vivo* proteolysis of abundant peptides ( $M_r < 3000$ ) in serum, generating heat maps of specimens *vs.* peptide IDs. The patterns were nota-

bly different, with some overlap, for sera from groups of breast, bladder, and prostate cancer patients. This comparison is a significant improvement over batch comparison of each tumor type separately to various control groups. They interpreted patterns to show ladders of cleavage products from particular peptides, probably due to progressive cleavage by aminopeptidases and/or carboxypeptidases. They identified fibrinopeptide A, C3f, and bradykinin as sources of these peptidome differences with potential relevance in cancer biology. Their presumption is that proteases persist in the sample (collected without protease inhibitors, of course) and continue to act *ex vivo*. If so, defined, standardized scaffolds might be more reliable substrates for such an assay of the unidentified proteases than the highly variable fibrin clot. What has long been treated as low molecular weight “noise” may be appropriate for further research on diagnostic potential [40]. In contrast, albumin-bound low molecular weight fragments of BRCA2, tyrosine kinases, other signaling molecules, and intracellular scaffolding proteins seem to reflect *in vivo* proteolytic cleavages [41]. One might prefer demonstration of *in vivo* proteolytic differences, as closer to the biology. This focus on peptides produced during the coagulation process requires serum specimens.

#### 2.4 Autoantibodies to tumor proteins

There is no *in vitro* protein counterpart to PCR for amplification of nucleic acids. However, the body has its own amplification system for proteins, namely an antibody response against epitopes on immunogenic tumor-associated proteins [42]. This approach utilizes the natural biological amplification of the immune response to generate immunoglobulin biomarkers for specific tumor immunogens. Obviously, these plasma or serum specimens should not be depleted of immunoglobulins. Hanash *et al.* pioneered the use of Western blots of 2-DE patterns from tumor cell lysates (A549) to test for immunoreactivity with serum from patients. Brichory *et al.* reported autoantibodies against annexin I and annexin II [43] and later against PGP 9.5 [44] in 30–50% of patients with adenocarcinomas of the lung, with high specificity. These assays have not yet been extended or expanded sufficiently to generate a panel for clinical evaluation, but the strategy is promising. The Western blots require MS analysis of the immunoreactive spots to determine what proteins can be identified, followed by independent methods to confirm the specificity of the autoantibody reactivity. Imafuku *et al.* [45] summarized autoantibody markers reported for various cancers. For example, for pancreatic cancers, autoantibodies have been reported against a truncated form of calreticulin [46], MUC1 [47], p53 [48], and Rad51 [49]. No one has yet put these and other candidate proteins together in a panel.

A complementary approach utilizes protein microarrays instead of 2-D gel Western blotting [50]. Tumor lysate fractions (A549 and H22 cell lines for lung cancer and LNCaP cell lines for prostate cancer) have been spotted robotically onto microarray slides after liquid-phase (IEF and HPLC)

fractionation. Haab has identified multiple spots with high reproducibility in clusters of related fractions. The patterns of immunoreactivity can be plotted on a heat map of fractions *vs.* patients and normals to show the differences and analyze for pattern differences between the groups for prostate cancers. The patterns were strikingly different for sera from patients with organ-confined prostate cancers *vs.* healthy controls compared with sera from patients with metastatic prostate cancer *vs.* the same healthy controls. Subsequent studies have employed MS of the immunoreactive spots to identify a protein or, often, a group of proteins that may account for the immunoreactivity. Otherwise, the results remain at the level of comparison of patterns between cancer patient sera and sera of normal controls or patients with other diseases [51].

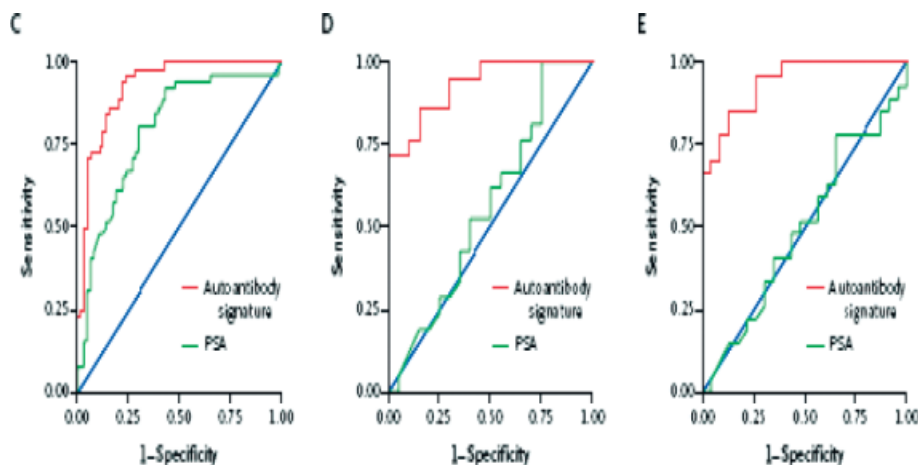
Phage-display methods are another practical method for screening for autoantibodies in cancer patients. Wang *et al.* [52] employed this method to generate a detector panel of 22 phase-peptides and then autoantibody signatures that gave remarkably good sensitivity and specificity for prostate cancer detection, far better than PSA (Fig. 4). As always, these findings need confirmation in larger clinical studies. Four of the 22 phage-peptides were confidently identified as proteins involved in transcription or translation: bromodomain-containing protein 2 (BRD2), eukaryotic translation initiation factor 4 (eIF4G1), and ribosomal proteins L13a and L22. These proteins are overexpressed in prostate cancer tumor specimens, by Western blots, and in the Oncomine gene expression database. Continuing work on this project is generating identifications for additional proteins from the 22 peptides. Confirmation by independent groups is the next step toward making this a useful test.

#### 2.5 Identifying proteins from protein pattern screens using ProteinChip SELDI-TOF-MS

Hundreds of papers have been published or presented using a variety of chip surfaces for fractionation of serum, followed directly by the SELDI method for MS. The first was a highly publicized paper by Petricoin *et al.* [53] about ovarian cancer. SELDI employs an artificial intelligence algorithm to compare groups of samples from cancer cases *versus* normals, generally finding that about 5 *m/z* peaks, among 15 000 or so from low molecular weight fragments, could discriminate the two groups with very high sensitivity and specificity. However, the results were seldom confirmed by independent groups, at least partly due to lack of reference standards and specimens, lack of standard operating procedures, and rapid changes in the MS technologies. The main proponents conducted these assays in batches without cross-comparison of results for different types of cancers. The NCI Early Detection Research Network invested significantly in this method over several years (<http://edrn.nci.nih.gov>).

When *m/z* peaks were analyzed further to identify the proteins responsible, high abundance host-response proteins, or variants of those proteins, were usually found. For





**Figure 4.** Receiver operator characteristics (ROC) curves based on multiplex analysis of autoantibody signatures with 22-phage-peptide detector panel, compared with PSA, for 60 prostate cancer patients and 68 controls. The red line is the phage-peptide-detector; the green line is the PSA test. The middle and right panels restrict the analysis to patients with PSA levels between 4 and 10 ng/mL, and 2.5 and 10 ng/mL, respectively, the critical zone for clinical decision-making. (Adapted from [52], with permission).

example, a thermostable variant of serum amyloid A1 was identified after analysis with the Ciphergen ProteinChip SELDI-TOF MS method in sera from 15 of 27 ovarian cancer patients, compared with 2 of 34 normal women. There was a double peak at 11.68 kDa and 11.52 kDa; the smaller peak is its N-terminal Arg-truncated isoform [54]. Incidentally, the heating step (10 min at 98°F) precipitated and removed the albumin.

Zhang *et al.* [55] reported a set of three discriminating peaks from a SELDI study of 153 patients with stage III invasive epithelial ovarian cancer, compared with 142 healthy women, 166 with benign pelvic masses, and 42 with other ovarian cancers. Results from two centers were confirmed with analyses of women from two other centers and then immunoassays of sera from a fifth center. MS showed the three peaks to be an increase of inter- $\alpha$  trypsin inhibitor heavy chain H4, decrease of a truncated form of transthyretin-10, and decrease of apolipoprotein A-1. These three marker candidates enhanced the sensitivity of CA-125 at a pre-determined specificity of 97% in patients with early-stage invasive ovarian cancers. So far, further validation and application of these findings have not been reported.

As the most lethal of cancers, pancreatic cancers are particularly cogent as a target for new diagnostic markers. From a SELDI-TOF analysis of 245 plasma samples from patients and controls [56], four  $m/z$  peaks were sufficient to discriminate pancreatic cancer from normals in the training and then testing format, with a claimed sensitivity of 91%, which increased to 100% when combined with CA-19-9, but with a false-positive rate of 19%, which would dominate any screening application. The authors rejected a peak probably due to serum amyloid A, but were not able to identify any of the four peaks reported. Bloomston *et al.* [57] utilized automated ProGEx 2-DE with 32 normal and 30 pancreatic cancer serum specimens; of 1744 protein spots, 154 were commonly overexpressed in cancer patient samples, of which 9 with just 4 identifications (two fibrinogen  $\gamma$  chain, complement factor B, and plasminogen) were reported as sufficient

to discriminate 30/30 cancer and 30/32 control sera. Fibrinogen  $\gamma$  may be important in the clinically observed hypercoagulability of the blood in such patients.

Similar results were obtained using 2-D gels in a study of pooled sera from 20 late-stage lung cancer patients and 20 healthy matched male donors in the Domont laboratory in Brazil [58]. Transthyretin monomer, haptoglobin- $\alpha$  2, and two isoforms of serum amyloid protein, as well as Ig  $\lambda$  chain, were up-regulated. A fragment of apolipoprotein A-1 was down-regulated in patients. Probably more systematic attention should be directed at these abundant proteins and their low-molecular weight conformers.

It has long been known that inflammatory or cellular stress responses somehow increase the concentrations of relatively abundant circulating proteins, so-called “acute-phase reactants”. Many conditions trigger increases in such proteins as haptoglobin,  $\alpha$ -1 anti-trypsin, serum amyloid A, transthyretin, and CRP. Nevertheless, in appropriate diagnostic situations, a rapid increase in concentrations or high measured level of these readily detected proteins may be helpful in guiding further diagnostic work-up for cancers, myocardial infarction, flares of rheumatoid arthritis, or other conditions.

PTMs may make these proteins more specific markers.

The most serious flaw in this blind-screening approach is the enormous number of comparisons tested, which should require a Bonferroni-type adjustment that would drastically reduce the statistical significance of the findings. Without identification of the proteins accounting for the  $m/z$  peaks, it is hard to interpret the results. Without reliable confirmation of the findings with additional sets of patients and in independent laboratories, the clinical value could not be demonstrated. There were two interlab comparison projects, one led by the NCI Early Detection Research Network [59] and one by the HUPO Plasma Proteome Project [60], which did show rather good correlation of  $m/z$  peak findings with the same reference specimens in several participating labs when complying with a common protocol.

## 2.6 Detection of differences in abundant proteins with antibody microarrays

Gao *et al.* [61] reported results from microarrays prepared by Haab with 84 antibodies in triplicate, against mostly abundant proteins. The antibody arrays were reacted with sera from 24 patients with adenocarcinoma of the lung, 32 patients with chronic obstructive lung disease, and 24 normal controls. Using a cut-off on immunoreactivity at which all 56 controls were non-reactive, the microarrays were able to identify 15 of the 24 lung cancer patients (62% sensitivity). As shown in Table 6, the most discriminating proteins were CRP (13-fold increase), serum amyloid A (2-fold increase), MUC1 (1.3–1.4-fold increase), and two isoforms of  $\alpha$ -1 antitrypsin (1.4-fold increase), while transferrin and gelsolin were decreased. We do not know whether the proteins were modified or intact.

Using a very similar set of 92 antibodies, the Haab laboratory probed sera from 61 pancreatic cancer patients, 31 patients with benign pancreatic disease, and 50 healthy controls in replicate experiments with the two-color, rolling circle amplification method. The proteins detected with reproducibly different binding levels between the patient classes reflected inflammation (high CRP,  $\alpha$ -1-antitrypsin, and serum amyloid A), immune response (high IgA), leakage of cell breakdown products (low plasma gelsolin), and possibly altered vitamin K usage or glucose regulation (high protein-induced vitamin K antagonist-II). Results were confirmed with immunoblot and antigen dilution experiments. A logistic-regression algorithm distinguished the cancer samples from the healthy control samples with 90% and 93% sensitivity and 90% and 94% specificity in duplicate experiments. The cancer samples were distinguished from the benign disease samples with 95% and 92% sensitivity and 88% and 74% specificity in duplicate experiment sets. The classification accuracies were significantly improved over those achieved using individual antibodies [62].

As noted above, numerous studies have reported acute-phase reactant proteins as potential biomarkers. Perhaps combinations of these should be evaluated in a multi-center study of several types of tumors.

## 3 Confirming and validating protein biomarker candidates and panels of biomarkers

The National Cancer Institute Early Detection Research Network set up a well-defined multi-stage scheme for discovery and validation of potential biomarkers several years ago. So far, few have made it past the discovery phase to replication in the same laboratory and then replication in a second laboratory. A much more complex stage of clinical validation then must follow, with appropriate screening scenarios to test and determine sensitivity, specificity, and positive predictive value in large populations. Guidance on proper epidemiological methodology and avoidance of bias and overfitting of results has been published recently by Ransohoff [63, 64] and Ioannidis [65]. Stratification of patients on clinical, pathological, or genetic criteria may be essential to account for the expected heterogeneity of mechanisms and, therefore, biomarkers in common cancers and other diseases. Therefore, sensitivity should not be expected to approach 100% for any one marker or even for a panel of very good markers. For screening purposes, the specificity must be as high as possible, to avoid false-positives, while still making the cut-off for quantitative differences appropriate for useful sensitivity as a screening test, avoiding many false-negatives, *i.e.*, missed diagnoses. The parameters for monitoring already diagnosed patients are quite different, since the diagnosis is already established, in contrast to a 4 in 10 000 incidence rate for ovarian cancer, for example, in adult female populations.

Due to heterogeneity of tumors and patients, and based on the extensive experience of recent years, it is unlikely that any single protein will be a highly predictive biomarker, especially in the screening scenario. More likely, a panel of several biomarkers will be required to achieve sufficient sensitivity to be useful for screening. The discovery of each of these candidates may come from different strategies. Optimization of such panels will be a complex challenge, seeking increasing sensitivity while keeping specificity high, to achieve useful predictive value for screening [63–65]. Multiplex ELISA and protein chip microarrays offer high throughput, but immunoassays in general may be incapable

**Table 6.** Antibody microarrays for abundant proteins differentiate lung cancer patients from controls (from [61])

Antibody	Fold difference in means		P-values from 1-way ANOVA	
	Tumor / Normal	Tumor / COPD	Tumor vs. Normal	Tumor vs. COPD
CRP	13.6	13.0	$1.1 \times 10^{-2}$	$2.2 \times 10^{-10}$
SAA	1.99	2.15	$1.8 \times 10^{-2}$	$1.7 \times 10^{-9}$
MUC1	1.30	1.42	$3.3 \times 10^{-3}$	$4.3 \times 10^{-5}$
AAT (1)	1.34	1.35	$9.1 \times 10^{-4}$	$3.1 \times 10^{-4}$
AAT (2)	1.42	1.33	$1.1 \times 10^{-4}$	$3.7 \times 10^{-3}$
Transferrin	0.73	0.71	$2.7 \times 10^{-4}$	$2.3 \times 10^{-3}$
Gelsolin	0.77	0.77	$5.8 \times 10^{-3}$	$4.7 \times 10^{-3}$

of differentiating size isoform variants. If no new epitope is formed, how will the immunoassay-based methods differentiate the variants? Automated combinations of immunoassays or chromatography with MS may provide such differentiation and identification. Instruments for 2-D LC-MS/MS that combine precise elution times with high mass accuracy will permit higher throughput identification of peptides and proteins using primarily LC-MS parameters.

Instead of “geographical” classification of cancers by organ of origin in the body, we will have molecular information with which to begin to classify tumors according to the mechanism that led to the tumor, which should be the basis for its selective treatment even before it can be detected with imaging. Combining the molecular marker with imaging techniques, perhaps utilizing nanoparticles, is a complementary path.

Leading journals and scientists have agreed that far more documentation is required for proteomics publications, both in the print version and the supplementary material, with guidelines now available [66] that should also inform the design of the whole study. There is also renewed attention at the National Cancer Institute, the National Institute for Standards and Technology, and various scientific societies to develop reference standards of peptides, proteins, and clinical specimens [67].

## 4 Comments

The need for reliable biomarkers for early diagnosis of various cancers remains compelling. Progress on prevention of clinical manifestations and on survival of those diagnosed with clinical cancers depends on such biomarkers. In some cases, like lung cancers, where there is a large-scale screening program with spiral CT scanning of the chest in 50 000 participants at increased risk due to cigarette smoking histories, molecular markers could be combined with imaging protocols to enhance clinical decision-making and improve cost-effectiveness of the total program.

The direct comparison of plasma or serum from patients and controls has not proved very productive in terms of reproducible results or practical diagnostic screening tests. Neither has the analysis of tumor specimens for mRNA or protein differences between cancers and normal or benign tissues. Part of the problem may be lack of attention to details of specimen collection, choice of controls, early replication, and exchange of reagents and specimens. Part of the challenge is setting specific criteria for the different kinds of assays for tumor-associated biomarkers in plasma, ranging from proteins secreted or released from the tumor to highly abundant acute-phase reactants and from normal proteins to proteins that may have specific and potentially diagnostic modifications. Several types of reference standards – peptides, proteins, and clinical specimens – will help calibrate instruments and quantitate proteins identified.

Furthermore, there is a greater appreciation now of the essential criteria for independent confirmation and then large-scale validation and of the necessity of anticipating heterogeneity among patients with a common diagnosis.

*I gratefully acknowledge research support from the Michigan Tri-Technology/Life Sciences Corridor, GR687, for the Proteomics Alliance for Cancer Research, and from the NIH/NLM 1 U54 DA021519–01A1 for the National Center for Integrative Biomedical Informatics; and 23XSIIOA for the NCI/SAIC Eastern Consortium for Proteomics Analysis of Mouse Models and Human Cancers.*

## 5 References

- [1] Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W. *et al.*, *Proteomics* 2005, 5, 13, 3226–3245.
- [2] Hood, L. E., Heath, J. R., Phelps, M. E., Lin, B., *Science* 2004, 306, 640–643.
- [3] Lin, B., White, J. T., Wei, L., Xie, T. *et al.*, *Cancer Res.* 2005, 8, 3081–3091.
- [4] Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M. *et al.*, *Science* 2005, 310, 644–648.
- [5] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, 1, 845–867.
- [6] Ludwig, J. A., Weinstein, J. N., *Nat. Rev. Cancer* 2005, 5, 845–856.
- [7] Rhodes, D. R., Yu, J., Shanker, K., Deshpande, N. *et al.*, *Neoplasia* 2004, 6, 1–6.
- [8] Hanahan, D., Weinberg, R. A., *Cell* 2000, 100, 57–70.
- [9] Rubin, M. A., Zhou, M., Dhanasekaran, S. M., Varambally, S. *et al.*, *JAMA* 2002, 287, 1662–1670.
- [10] Varambally, S., Dhanasekaran, S. M., Barrette, T. R., Sanda, M. G. *et al.*, *Nature* 2002, 419, 624–629.
- [11] Beer, D. G., Kardia, S. L., Huang, C. C., Giordano, T. J. *et al.*, *Nat. Med.* 2002, 8, 816–824.
- [12] Chen, G., Gharib, T. G., Wang, H., Huang, C. C. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, 98, 9824–9829.
- [13] Varambally, S., Yu, J., Laxman, B., Rhodes, D. R. *et al.*, *Cancer Cell* 2005, 8, 393–406.
- [14] Xiao, T., Ying, W., Li, L., Hu, H. *et al.*, *Mol. Cell. Proteomics* 2005, 4, 10, 1480–1486.
- [15] Thongboonkerd, V. (Ed.), *Proteomics of Human Body Fluids: Principles, Methods, and Applications*. Humana Press, Totowa 2006, in press.
- [16] Jurgens, M., Appel, A., Heine, G., Neitz, S. *et al.*, *Comb. Chem. High Throughput Screen.* 2005, 8, 757–765.
- [17] Gao, W-M, Chadha, M. S., Berger, R. P., Omenn, G. S. *et al.*, *J. Neurotrauma* 2006 (in press).
- [18] Tyan, Y-C, Wu, H-Y, Lai, W-W, Su, W-C *et al.*, *J. Proteome Res.* 2005, 4, 1274–1286.
- [19] Celis, J. E., Gromov, P., Cabezon, T., Moreira, J. M. A. *et al.*, *Mol. Cell. Proteomics* 2004, 3, 327–344.
- [20] Dua, R. S., Isacke, C. M., Gui, G. P., *J. Clin. Oncol.* 2006, 24, 1209–1216.

- [21] Utleg, A. G., Yi, E. C., Xie, T., Shannon, P. *et al.*, *Prostate* 2003, **56**, 150–161.
- [22] Thiery, J. P., *Nat. Rev. Cancer* 2002, **2**, 442–454.
- [23] Keshamouni, V. G., Michailidis, G., Grasso, C. S., Anthwal, S. *et al.*, *J. Proteome Res.* 2006, online.
- [24] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B. *et al.*, *Mol. Cell. Proteomics* 2004, **3**, 1154–1169.
- [25] Gu, S., Liu, Z., Pan, S., Jiang, Z. *et al.*, *Mol. Cell. Proteomics* 2004, **3**, 998–1008.
- [26] Witt, A. E., Hines, L. M., Collins, N. L., Hu, Y. *et al.*, *J. Proteome Res.* 2006, **5**, 599–610.
- [27] Nishizuka, S., Charboneau, L., Young, L., Major, S. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, **25**, 14229–14234.
- [28] Nishizuka, S., Chen, S. T., Gwadry, F. G., Alexander, J. *et al.*, *Cancer Res.* 2003, **63**, 5243–5250.
- [29] Nettikadan, S., Radke, K., Johnson, J., Xu, J. *et al.*, *Mol. Cell. Proteomics* online 19 Feb 2006, M500350–MCP200.
- [30] Irish, J. M., Kotecha, N., Nolan, G. P., *Nat. Rev. Cancer* 2006, **6**, 146–154.
- [31] Hanash, S., *Nat. Rev. Cancer* 2004, **4**, 638–644.
- [32] Anderson, L., Hunter, C. L., *Mol. Cell. Proteomics* 2006, **5**, 573–588.
- [33] Joos, T. O., Stoll, D., Templin, M. F., *Curr. Opin. Chem. Biol.* 2002, **6**, 76–80.
- [34] Haab, B., *Mol. Cell. Proteomics* 2005, **4**, 377–383.
- [35] Kuster, B., Schirle, M., Mallick, P., Aebersold, R., *Nat. Rev. Mol. Cell Biol.* 2005, **6**, 577–583.
- [36] Craig, R., Cortens, J. P., Beavis, R. C., *Rapid Commun. Mass Spectrom.* 2005, **19**, 1844–1850.
- [37] Yang, Z., Hancock, W. S., Chew, T. R., Bonilla, L., *Proteomics* 2005, **5**, 3353–3366.
- [38] Petricoin, E. F., Ornstein, D. K., Liotta, L. A., *Urol. Oncol.* 2004, **22**, 322–328.
- [39] Villanueva, J., Shaffer, D. R., Philip, J., Chaparro, C. A. *et al.*, *J. Clin. Invest.* 2006, **116**, 271–284.
- [40] Liotta, L. A., Petricoin, E. F., *J. Clin. Invest.* 2006, **116**, 26–30.
- [41] Lowenthal, M. S., Mehta, A. I., Frogale, K., Bandle, R. W. *et al.*, *Clin. Chem.* 2005, **51**, 1933–1945.
- [42] Hanash, S., *Nat. Biotechnol.* 2003, **21**, 37–38.
- [43] Brichory, F., Misek, D. E., Yim, A. M. *et al.*, *Proc. Natl. Acad. Sci. USA* 2001, **98**, 9824–9829.
- [44] Brichory, F., Beer, D., Le Naour, F., Giordano, T. *et al.*, *Cancer Res.* 2001, **61**, 7908–7912.
- [45] Imafuku, Y., Omenn, G. S., Hanash, S., *Dis. Markers* 2004, **20**, 149–153.
- [46] Hong, S. H., Misek, D. E., Wang, H., Puravs, E. *et al.*, *Cancer Res.* 2004, **64**, 5504–5510.
- [47] Hamanaka, Y., Suehiro, Y., Fukui, M., Shikichi, K. *et al.*, *Int. J. Cancer* 2003, **103**, 97–100.
- [48] Raedle, J., Oremek, G., Welker, M., Roth, W. K. *et al.*, *Pancreas* 1996, **13**, 241–246.
- [49] Maacke, H., Hundertmark, C., Miska, S., Voss, M. *et al.*, *J. Cancer Res. Clin. Oncol.* 2002, **128**, 219–222.
- [50] Bouwman, K., Qiu, J., Zhou, H., Schotanus, M. *et al.*, *Proteomics* 2003, **3**, 2200–2207.
- [51] Qiu, J., Madoz-Gurpide, J., Misek, D. E., Kuick, R. *et al.*, *J. Proteome Res.* 2004, **3**, 261–267.
- [52] Wang, X., Yu, J., Sreekumar, A., Varambally, S. *et al.*, *N. Engl. J. Med.* 2005, **353**, 1224–1235.
- [53] Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J. *et al.*, *Lancet* 2002, **359**, 572–577.
- [54] Moshkovskii, S. A., Serebryakova, M. V., Kutteykin-Teplyakov, K. B., Tikhonova, O. V. *et al.*, *Proteomics* 2005, **5**, 3790–3797.
- [55] Zhang, Z., Bast, R. C. Jr., Yu, Y., Li, J. *et al.*, *Cancer Res.* 2004, **64**, 5882–5890.
- [56] Honda, K., Hayashida, Y., Umaki, T., Okusaka, T. *et al.*, *Cancer Res.* 2005, **65**, 10613–10622.
- [57] Bloomston, M., Zhou, J. X., Rosemurgy, A. S., Frankel, W. *et al.*, *Cancer Res.* 2006, **66**, 2592–2599.
- [58] Maciel, C. M., Junqueira, M., Paschoal, M. E., Kawamura, M. T. *et al.*, *J. Exp. Ther. Oncol.* 2005, **5**, 31–38.
- [59] Semmes, O. J., Feng, Z., Adam, B. L., Banex, L. L. *et al.*, *Clin. Chem.* 2005, **51**, 102–112.
- [60] Rai, A. J., Stemmer, P. M., Zhang, Z., Adam, B.-L. *et al.*, *Proteomics* 2005, **5**, 3467–3474.
- [61] Gao, W. M., Kuick, R., Orzechowski, R. P., Misek, D. E. *et al.*, *BMC Cancer* 2005, **5**, 110 (<http://www.biomedcentral.com/1471-2407/5/110>).
- [62] Orzechowski, R., Hamelinck, D., Li, L., Gliwa, E. *et al.*, *Cancer Res.* 2005, **65**, 11193–11202.
- [63] Ransohoff, D. E., *J. Natl. Cancer Inst.* 2005, **97**, 315–319.
- [64] Ransohoff, D. E., *Nat. Rev. Cancer* 2005, **5**, 142–149.
- [65] Ioannides, J. P. A., *PLoS Med.* 2005, **2**, e124.
- [66] Carr, S., Aebersold, R., Baldwin, M., Burlingame, A. *et al.*, *Mol. Cell. Proteomics* 2004, **3**, 351–353.
- [67] Barker, P. E., Wagner, P. D., Stein, S. E., Bunk, D. M. *et al.*, *Clin. Chem.* 2006, **52**, 1669–1674.