Mass Spectrometry: Uncovering the Cancer Proteome for Diagnostics

Da-Elene van der Merwe,[*] Katerina Oikonomopoulou,[*,†] John Marshall,[‡] and Eleftherios P. Diamandis[*,†]

[*]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario M5G1X5, Canada;
[†]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario M5G1L5, Canada;
[‡]Department of Chemistry and Biology, Ryerson University, Toronto, Ontario M5G1G3, Canada

I. Current Cancer Biomarkers
II. Early Detection
  A. When Is an Early Detection Program Warranted?
III. The Need for New Diagnostic Strategies
IV. Mass Spectrometry
  A. Ionization Source
  B. Mass Analyzers
  C. Protein Identification
  D. Quantitation
V. Mass Spectrometry-Based Diagnostics
  A. Mass Spectrometry as a Tissue Imaging Tool
  B. Mass Spectrometry as a Biomarker Discovery Tool
  C. Mass Spectrometry as a Cancer Diagnostic Tool
VI. Current Limitations of Diagnostic Mass Spectrometry
  A. Preanalytical
  B. Analytical
  C. Postanalytical
VII. Suggestions for Future Progress
VIII. Future Direction
References

Despite impressive scientific achievements over the past few decades, cancer is still a leading cause of death. One of the major reasons is that most cancer patients are diagnosed with advanced disease. This is clearly illustrated with ovarian cancer in which the overall 5-year survival rates are only 20–30%. Conversely, when ovarian cancer is detected early (stage 1), the 5-year survival rate increases to 95%. Biomarkers, as tools for preclinical detection of cancer, have the potential to revolutionize the field of clinical diagnostics. The emerging field of clinical proteomics has found applications across a wide spectrum of cancer research. This chapter will focus on mass spectrometry as a proteomic technology implemented in three areas of cancer: diagnostics, tissue imaging, and biomarker discovery. Despite its power, it is also important to realize the preanalytical,

Advances in CANCER RESEARCH
0065-230X/07 $35.00
DOI: 10.1016/S0065-230X(06)96002-3
Copyright 2007, Elsevier Inc. All rights reserved.

analytical, and postanalytical limitations currently associated with this methodology. The ultimate endpoint of clinical proteomics is individualized therapy. It is essential that research groups, the industry, and physicians collaborate to conduct large prospective, multicenter clinical trials to validate and standardize this technology, for it to have real clinical impact. © 2007 Elsevier Inc.

I. CURRENT CANCER BIOMARKERS

Currently, hundreds of tumor markers exist, yet most of them fall short of expectation. Clinicians expect that a marker should be beneficial to their patients in terms of improved morbidity, mortality, and quality of life. To illustrate the point, even if a biomarker is able to detect relapse a few months prior to clinical symptoms, if effective treatment does not exist, this information does not necessarily translate into improved outcome. Moreover, knowledge of tumor marker elevation may be potentially harmful since it shortens disease-free survival and adds to patient anxiety.

Despite their known shortcomings, tumor markers continue to be used in a variety of clinical settings. Some of the current applications of tumor markers and their limitations are listed in Table I.

Currently, controversy exists regarding the optimal use of tumor markers among clinicians and laboratory medicine specialists. This is reflected in practice guidelines developed by various professional societies. In 1998, the National Academy of Clinical Biochemistry (NACB)[1] sponsored a consensus conference to develop guidelines for the analytical performance and clinical utility of tumor markers (Fleisher *et al.*, 2002). The recommendations focused on pre- and postanalytical concerns, the use of reference intervals, and the manner in which tumor markers should be used clinically, with specific attention to screening, diagnosis, monitoring, or prognosis.

In the mid-1980s, a working group consisting of German scientists, physicians, and representatives of the diagnostics industry were established. In 1993, this group published a consensus statement on the criteria for use of tumor markers with respect to clinical relevance, analytical methods, and manufacturing requirements (European Group on Tumour Markers, 1999;

[1] Abbreviations: NACB, National Academy of Clinical Biochemistry; EGTM, European Group on Tumor Makers; SEER, Surveillance, epidemiology, and end results; WHO, World Health Organization; EDRN, Early Detection Research Network; ELISA, Enzyme-linked immunosorbent assay; CGAP, Cancer Genome Anatomy Project; SAGE, Serial analysis of gene expression; EST, Expressed sequence tag; SELDI-TOF, Surface-enhanced laser desorption/ ionization-time-of-flight; MUDPIT, Multidimentsional protein identification technology; HPLC, High-performance liquid chromatography; ESI, Electrospray ionization; MALDI, Matrix-assisted laser desorption; FT-ICR MS, Fourier transform ion cyclotron ionization resonance mass spectrometer; CID, Collisional-induced dissociation; SILAC, Stable-isotope labeling with amino acids in cell culture; ICAT, Isotope-coded affinity tag; LCM, Laser capture microdissection; IMAC, Immobiliszed metal affinity capture; NCI, National Cancer Institute.

Table I Current Applications of Tumor Markers

Application	Clinical value	Comments
1. Population screening	Limited	Low diagnostic sensitivity and specificity
2. Diagnosis	Limited	Low diagnostic sensitivity and specificity
3. Prognosis	Limited	Not sufficiently accurate
4. Tumor staging	Limited	Not sufficiently accurate
5. Tumor localization and targeted therapy	Limited	Low specificity, low efficiency
6. Detection of recurrence	Controversial	Short lead time, unavailable effective therapy, misleading information due to low specificity
7. Monitoring therapeutic response	Important	Biomarker usually superior to imaging modalities
8. Prediction of therapeutic response	Important	Therapy given only to those who will benefit sparing others from toxic side effects

Van Dalen, 1993). This group was formally constituted as the European Group on Tumor Markers (EGTM) in 1997. Many other clinical organizations, such as the American Society of Clinical Oncology (ASCO), formulated their own recommendations [Bast *et al.*, 2001a,b; Tumor Marker Expert Panel (ASCO), 1996]. Reviews on practice guidelines for tumor markers have been published (Duffy *et al.*, 2003; Loi *et al.*, 2004; Sturgeon, 2001, 2002) as well as strategies for their development (Oosterhuis *et al.*, 2004).

II. EARLY DETECTION

Cancer continues to be diagnosed late, when therapeutic options are limited to palliative care. In our battle against cancer, emphasis should shift from clinical diagnosis to preclinical disease detection, before cancer metastasizes and becomes incurable. In an era of evidence- and outcomes-based medicine, the following questions are relevant: (1) *Why do we need early cancer detection?* and (2) *When is an early disease detection program warranted?* (Etzioni *et al.*, 2003).

The answer to the first question is twofold: (1) Treatment of advanced disease is almost never curative. This is illustrated in the very modest gains in survival rates of patients diagnosed with advanced cancers of different organs from 1973 to 1997 (National Cancer Institute, 2002). (2) Early detection of cancer improves outcome. Ovarian cancer is a good example

Table II Projected Changes in Survival with Early Detection[a]

Cancer site	Tumors localized when detected (%)	Five-year survival rate (%)	Five-year survival rate if all tumors were localized when detected (%)
Lung	19	16	49
Colorectal	41	64	90
Breast	65	87	97
Prostate	65	90	100

[a]Based on data from SEER (National Cancer Institute, 2002) for cases diagnosed between 1990 and 1999 inclusive. Cases with *in situ* or unstaged disease have been excluded. The favorable overall 5-year survival among breast and prostate cancer patients is partly due to the prevalence of screening during the calendar years considered. Reprinted from Etzioni *et al.* (2003) with permission from copyright owners.

where early detection can have a major impact. More than two-thirds of ovarian cancer cases are detected at an advanced stage, when the cancer cells have spread away from the ovarian surface and have disseminated throughout the peritoneal cavity (Menon and Jacobs, 2002; Meyer and Rustin, 2000). The resulting 5-year survival rate is 20–30% with the best available treatment. Conversely, when the cancer is detected early (stage 1), conventional therapy leads to 95% 5-year survival (Bast *et al.*, 1983; Cohen *et al.*, 2001; Jacobs *et al.*, 1999; Menon and Jacobs, 2000). Similar figures apply to colon and other cancers. The best evidence comes from the Surveillance Epidemiology and End Results (SEER) program conducted by the National Cancer Institute (2002). Survival is excellent for the main cancers when early-stage disease is treated with existing therapies (Table II).

A. When Is an Early Detection Program Warranted?

According to the World Health Organization (WHO) the following criteria need to be fulfilled: (1) the disease must be common and associated with serious morbidity and mortality, (2) screening tests must be able to accurately detect early-stage and potentially curable disease, (3) treatment after detection through screening must show a significant advantage relative to the treatment without screening, and (4) evidence that the overall potential benefits outweigh the potential harms and costs of screening (Winawer *et al.*, 1995). For early detection to be an effective and practical approach, screening tests must satisfy four basic requirements. (1) Screening tests should distinguish healthy individuals from cancer cases with a high degree of accuracy, that is, high sensitivity and specificity and high positive and

negative predictive values. (2) Detection should be possible before the disease progresses to an advanced stage, when treatment is less effective. (3) Screening tests should ideally differentiate between aggressive lesions (which require treatment) and benign tumors, avoiding the problem of overdiagnosis. (4) Tests should be inexpensive, minimally invasive, and well accepted by the targeted population.

Although screening tests are currently in use for some cancers, very few satisfy these requirements.

III. THE NEED FOR NEW DIAGNOSTIC STRATEGIES

Refinements in more conventional diagnostic strategies, such as imaging, have had a substantial benefit to patients over the last 25 years. The potential to detect early breast cancer by mammography or the ability of computed tomography, ultrasonography, and magnetic resonance imaging to reveal small masses or tumor metastasis are but a few examples. However, hybrid strategies, combining imaging with other modalities should work better. Novel biomarkers, as additional tools to detect preclinical cancers, have the potential to revolutionize the way we diagnose and manage cancer in the future.

The rapidly expanding field of cancer biomarker discovery prompted the establishment of the Early Detection Research Network (EDRN) by the National Cancer Institute (NCI) (Srivastava and Kramer, 2000). The purpose of the EDRN is to coordinate research among biomarker development laboratories, biomarker validation laboratories, clinical repositories, and population screening programs with the hope to facilitate collaboration and to promote efficiency and rigor in research. The objectives of the EDRN for biomarker development and validation can be summarized in five consecutive phases: (1) preclinical exploratory, (2) clinical assay and validation, (3) retrospective longitudinal, (4) prospective screening, and (5) cancer control (Sullivan *et al.*, 2001).

Until recently, biomarker discovery was a laborious, linear, and slow process, where each candidate biomarker is first identified and then validated for specificity and sensitivity by using mainly an enzyme-linked immunosorbent assay (ELISA). With the advent of the post-genomic era, powerful new approaches are being realized. One approach is to use bioinformatics such as digital differential display and *in silico* Northern analysis utilizing SAGE, EST, cDNA arrays, or other parallel (Brenner and Johnson, 2000) nucleic acid analysis techniques, and the databases of the Cancer Genome Anatomy Project (CGAP) to compare gene expression between healthy and cancerous tissues in order to identify overexpressed

genes (Hermeking, 2003; Hess, 2003; Polyak and Riggens, 2001; Tuteja and Tuteja, 2004; Yousef *et al.*, 2003). Gene expression analysis by microarray technology is another method that identifies overexpressed genes in cancer, with the potential to develop cancer biomarkers (Hampton and Frierson, 2003; Hellstrom *et al.*, 2003; Lu *et al.*, 2004; Welsh *et al.*, 2001, 2003; Zarrinkar *et al.*, 2001). However, some of the best cancer biomarkers (such as PSA) are not overexpressed in cancer (Magklara *et al.*, 2000).

The emerging field of clinical proteomics is not only well suited to the discovery and implementation of new biomarkers, but it could also be applied across the spectrum of cancer research (Fig. 1). Proteomics refers to the systematic study of the total protein complement (proteome) encoded and expressed by a genome or by a particular cell, tissue, or organism (Pusch *et al.*, 2003). Many researchers have hypothesized that the best cancer biomarkers will likely be secreted proteins (Welsh *et al.*, 2003). Approximately 20–25% of all cell proteins are secreted. Proteins, or their fragments, originating from cancer cells or their microenvironment, may eventually enter the circulation. The patterns of expression of these proteins could be analyzed by mass spectrometry in combination with mathematical algorithms. Proteomic pattern diagnostics include proteomic

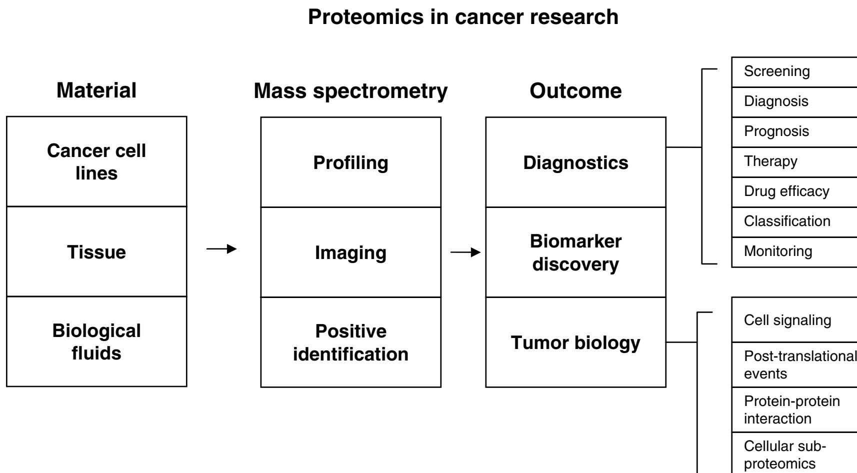

Fig. 1 Application of clinical proteomics in cancer research. Clinical material (cell lines, tissues, or biological fluids) is analyzed by mass spectroscopy with or without chromatographic separation for either imaging, proteomic profiling, or for identification of putative biomarkers. This analysis can lead to development of novel diagnostics or for understanding tumor biology.

pattern profiling of serum by surface-enhanced laser desorption/ionization-time-of-flight (SELDI-TOF) mass spectrometry combined with bioinformatic tools (Petricoin *et al.*, 2002c). The rest of this chapter will focus on mass spectrometry as a tool for biomarker discovery and as a diagnostic platform for cancer.

IV. MASS SPECTROMETRY

The concept of global protein analysis as a complete inventory of human proteins was proposed 20 years ago (Anderson and Anderson, 2002; Anderson *et al.*, 2004), and proteomic research was driven in the mid-1990s by the development in three areas: two-dimensional gel electrophoresis, mass spectrometry, and bioinformatic databases. In "top-down" proteomics, intact proteins are analyzed. In "bottom-up" proteomics, the proteins are proteolytically cleaved intentionally, using enzymes. In contrast, the endogenous peptides of serum or plasma presumably result from physiological proteolysis *in vivo* or *in vitro*.

Commonly available mass spectrometers are sensitive to the hundreds of atto mols and zeptomolar sensitivity has been demonstrated (Dick Smith). However, in practice sensitivity is overwhelmingly dependent on sample preparations.

Mass spectrometry-based proteomics has become the method of choice for the analysis of complex protein samples. Mass spectrometers have been used for many decades as diagnostic tools in clinical laboratories and have enjoyed many successes in the area of identification and quantification of relatively small molecules (molecular mass < 1000 Da). Recent interest in this technology for studying larger molecules, such as nucleic acids and proteins, has escalated significantly (Aebersold and Mann, 2003; Fenn *et al.*, 1989; Pedrioli *et al.*, 2004; Tyers and Mann, 2003). This has been made possible not only by the availability of genome sequence databases, but particularly by the discovery and development of novel protein ionization methods recognized by the 2002 Nobel Prize in chemistry.

A mass spectrometer consists of an ion source, a mass analyzer that measures the mass-to-charge ratio (m/z) of the ionized analytes and a detector that registers the number of ions at each m/z value (Aebersold and Mann, 2003). A typical proteomic experiment usually consists of five stages. (1) The proteins to be analyzed, present in cell lysates, tissues, or fluids are separated by various fractionation or affinity selection techniques (Lim and Elenitoba-Johnson, 2004). This defines the "subproteome" to be analyzed. The most powerful recent strategy integrates different separation methods as multidimensional combinations (MUDPIT) such as ion-exchange

and reverse-phase HPLC. (2) Enzymatic protein degradation to peptides (usually by trypsin). MS of whole proteins is less sensitive than peptide MS and the intact protein by itself may not be as easily detected, although methods for examining large proteins are rapidly advancing. (3) Peptides are routinely separated by high-performance liquid chromatography in very fine capillaries and eluted into an electrospray ion source where they are nebulized in small, highly charged droplets. After evaporation, multiple protonated peptides enter the mass spectrometer. (4) A mass spectrum of the peptides eluting at each time point is taken. (5) These peptides are prioritized for fragmentation and a series of tandem mass spectrometric (MS/MS) experiments ensues to obtain sequence information. Identified peptides are matched against protein sequence databases to eventually identify the proteins of interest.

Essential to proteomic studies is the simplification of a complex mixture of proteins into less complex components. In general, measurement of peptide masses by MS is experimentally and mathematically (Mann *et al.*, 2001) simpler than the calculation of intact protein masses. The ability to accurately determine the mass of a unique peptide that originates from a particular protein greatly facilitates the identification of that protein (Hunt and Shabanowitz, 1987; Smith and Anderson, 2002).

A. Ionization Source

Mass spectrometric measurements are carried out in the gas phase on ionized analytes. Two techniques are most commonly used to volatize or ionize the proteins or peptides, namely electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp, 1988; Nakanishi *et al.*, 1994). ESI ionizes analytes out of a solution and is therefore readily coupled to liquid-based separation tools (e.g., chromatographic and electrophoretic). MALDI ionizes the samples out of dry, crystalline matrix via laser pulses. MALDI-MS is normally applied to relatively simple peptide mixtures, compared to ESI-MS combined with liquid-chromatography (LC-MS), which is preferred for the analysis of complex samples. A variant MALDI technology, which has been used extensively in diagnostics, is surface-enhanced laser desorption ionization (SELDI) (Merchant and Weinberger, 2000). In this technology, a surface (ProteinChip^{TM}) functions as a solid phase extraction tool. The objective is to overcome the requirement for purification and separation of proteins prior to MS analysis (Aebersold and Goodlett, 2001).

B. Mass Analyzers

The mass analyzer separates ions according to *m/z* ratio. In terms of proteomics, its key parameters are sensitivity, resolution, mass accuracy, and the ability to generate information-rich mass spectra from peptide fragments (Mann *et al.*, 2001; Pandey and Mann, 2000; Wilkins *et al.*, 1998). Four basic types of mass analyzers are commonly used in proteomic research: the ion trap, time-of-flight (TOF), quadruple, and Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) analyzer. They all have different characteristics and can be used on their own or in combination with each other to optimize results (Lim and Elenitoba-Johnson, 2004).

C. Protein Identification

1. PEPTIDE MASS FINGERPRINTING

This is the simplest method for protein identification which combines enzymatic digestion, mass spectrometry, and data analysis. The peptides generated are analyzed by MS and the masses are compared with theoretical mass spectra of proteins listed in databases. Software algorithms for peptide mass mapping include PeptIdent/MultiIdent and ProFound (MacCoss *et al.*, 2002; Zhang and Chait, 2000).

2. PEPTIDE SEQUENCING BY TANDEM MASS SPECTROMETRY

This technique is based on collisional-induced dissociation (CID) that randomly cleaves peptide bonds between adjacent amino acid residues. This yields ion series that eventually reveal the amino acid sequence of a peptide.

D. Quantitation

Small molecules are routinely quantified on triple stage quadrupole mass spectrometers and this may one day be extended to peptides. A quantitative dimension has been added to MS experiments by stable-isotope dilution (SILAC), which is based on the principle that pairs of chemically identical analytes of different stable-isotope composition can be differentiated by

MS owing to their mass difference, and that the ratio of signal intensities for such analyte pairs accurately indicates the abundance ratio for the two analytes (Conrads *et al.*, 2002; Mirgorodskaya *et al.*, 2000; Yao *et al.*, 2001). Another technology, isotope-coded affinity tags (ICAT), relies on stable isotope labeling of cysteine residues (Gygi *et al.*, 1999; Von Haller *et al.*, 2003a, b). The advantage of this method is that it allows evaluation of low-abundance proteins and proteins at both extremes of molecular weight and isoelectric point. Absolute quantitation requires prior identification of the analyte and the use of external or internal standards.

V. MASS SPECTROMETRY-BASED DIAGNOSTICS

Mass spectrometry has been used in two different settings in the area of cancer diagnostics, first for the discovery of novel cancer biomarkers and second as a cancer diagnostic and imaging tool. The discovery of biomarkers and their use as early detectors of cancer is based on the hypothesis that a complex interplay exists between a tumor and its host microenvironment (Liotta and Kohn, 2001). As blood perfuses through a diseased organ, the serum protein profile is altered as a result of ongoing physiological and pathological events. This may include proteins being overexpressed and/or abnormally shed, clipped, modified, or removed due to abnormal activation of the proteolytic degradation pathway, generating a unique signature in blood (Fig. 2). As a consequence, the expressed serum protein profile is different between normal and diseased states. This creates a unique opportunity to exploit accessible body fluids, such as serum, urine, saliva, seminal plasma, malignant ascites, or cerebrospinal fluid, for the discovery of novel biomarkers.

A. Mass Spectrometry as a Tissue Imaging Tool

A recent advance, laser capture microdissection (LCM) provides a means of rapidly procuring pure cell populations from the surrounding heterogeneous tissue, allowing the use of tissue as an additional medium to discover novel biomarkers (Banks, 1999; Emmert-Buck *et al.*, 1996). The concept of MALDI-MS-based imaging mass spectrometry was introduced in 1997 by Caprioli *et al.* (1997). MS is used to map the distribution of peptides and proteins directly from thin tissue sections and allows visualization of 500–1000 individual protein signals in the molecular weight range from 2000 to 200,000. Matrix is deposited uniformly over the section and

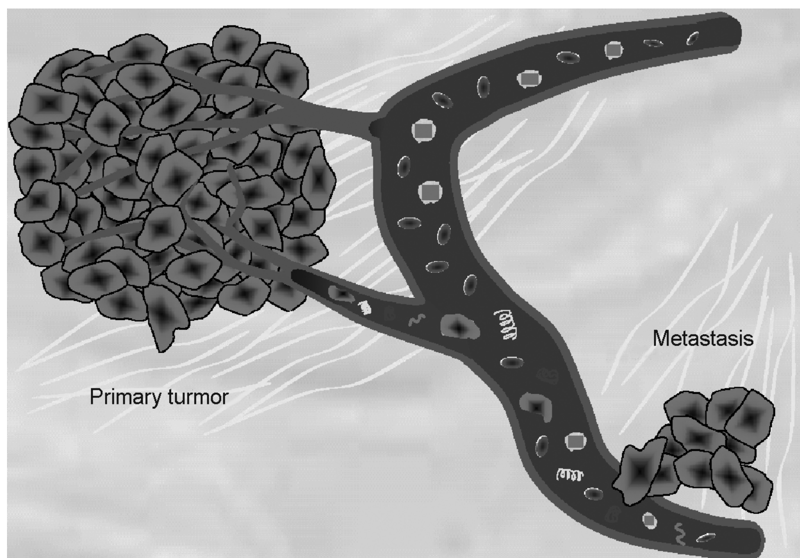

Fig. 2 Schematic representation of blood composition after contribution of molecules or cells by tumor due to angiogenesis or tissue destruction. The enrichment of blood with tumor- or microenvironment-derived components can be used for diagnostics.

analysis of the tissue is performed over a predetermined two-dimensional array or grid generating a full mass spectrum at each grid coordinate. Each spectrum is generated with an average of 15–50 consecutive laser shots at each coordinate. From the intensity of a given *m/z* value, a density map or image can be constructed. It is essential to maintain three conditions: (1) the deposition process must not disperse or translocate proteins within the section, (2) the matrix solution must wet the tissue surface to form crystals which contain cocrystallized proteins, and (3) the crystal dimensions must be smaller than the image resolution (Chaurand *et al.*, 2002). Imaging mass spectrometry is still in an early developmental stage and many improvements in sample preparation, handling, and instrumentation can be expected in future. However, this technique yields a wealth of information about the protein pattern trends within a tissue sample, and differentially expressed protein profiling between healthy and cancerous tissues has already been explored for novel cancer biomarker identification (Schwartz *et al.*, 2004; Yanagisawa *et al.*, 2003). More recently, this method has been used to predict tumor response to molecular therapeutics (Reyzer *et al.*, 2004). This may become an important means to delineate surgical margins in real time during surgery.

B. Mass Spectrometry as a Biomarker Discovery Tool

The use of MS as a biomarker discovery technique is conceptually straightforward. Fluids or tissue extracts from a diseased group, as well as a control group, are analyzed by MS and the differentially expressed peaks are identified. These peaks potentially represent molecules that could be measured with simpler and cheaper techniques, such as ELISA, for the purpose of diagnosis and management of cancer. A list of candidate biomarkers identified by MS is shown in Table III (Cho *et al.*, 2002; Johnson *et al.*, 1994; Koomen *et al.*, 2005; Zhang *et al.*, 2004). So far, MS discovery efforts have focused on three subsets of the proteome: (1) polypeptides and whole

Table III Serum Concentration of Some Abundant Proteins, Classical Cancer Biomarkers, and Putative New Cancer Biomarkers Identified by Mass Spectrometry[a]

Compound	Approximate concentration (pmol/liter)	Biomarker for cancer type	References
Serum proteins			
Albumin	600,000,000	–	Johnson *et al.*, 1994
Immunoglobulins	30,000,000	–	Johnson *et al.*, 1994
C-reactive protein	40,000	–	Johnson *et al.*, 1994
Classical tumor markers			
α-Fetoprotein	150	Hepatoma, testicular	Johnson *et al.*, 1994
Prostate-specific antigen	140	Prostate	Johnson *et al.*, 1994
Carcinoembryonic antigen	30	Colon, pancreas, lung, breast	Johnson *et al.*, 1994
Choriogonadotropin	20	Testicular, choriocarcinoma	Johnson *et al.*, 1994
β-Subunit of choriogonadotropin	2	Testicular, choriocarcinoma	Johnson *et al.*, 1994
Mass spectrometry-identified proteins			
Apolipoprotein A1	40,000,000	Ovarian, pancreatic	Liotta *et al.*, 2003; Zhang *et al.*, 2004
Transthyretin fragment	6,000,000	Ovarian	Zhang *et al.*, 2004
Inter-α-trypsin inhibitor fragment	4,000,000	Ovarian, pancreatic	Koomen *et al.*, 2005; Zhang *et al.*, 2004
Haptoglobin-α-subunit	1,000,000	Ovarian, pancreatic	Koomen *et al.*, 2005
Vitamin D-binding protein	10,000,000	Prostate	Zhang *et al.*, 2004
Serum amyloid A	20,000,000	Nasopharyngeal, pancreatic	Koomen *et al.*, 2005; Cho *et al.*, 2002
α_1-Antitrypsin	10,000,000	Pancreatic	Koomen *et al.*, 2005
α_1-Antichymotrypsin	5,000,000	Pancreatic	Koomen *et al.*, 2005

[a]Reproduced from Diamandis and van der Merwe (2005) with permission from copyright owners.

proteins that can be analyzed by electrophoresis with or without prior fractionation, (2) enzymatic peptide fragments separated by liquid chromatography and analyzed with either ESI or MALDI, typically after one or more chromatographic or other fractionation steps, and (3) naturally occurring peptides (the peptidome), which provide a complementary picture of many events at the low mass end of the plasma proteome (Liotta *et al.*, 2003; Loboda and Krutchinsky, 2000; Marshall and Jankowski, 2004; Villanueva *et al.*, 2004). The biggest challenge in uncovering potential biomarkers present in serum lies in the complexity and dynamic range of the proteome. Various prefractionation steps have been applied to mine into the subproteome in order to reach the low-abundance and likely the most informative molecules (Fig. 3).

C. Mass Spectrometry as a Cancer Diagnostic Tool

The concept and utility of multivariate protein markers as opposed to a single indicator to diagnose disease has been established for some time. More than 20 years ago, it became clear that different tumor cell types could be distinguished based on patterns of metabolites analyzed by GC-MS (Jellum *et al.*, 1981). Investigators are currently using two types of proteomic technologies to mine the proteomic signature in order to differentiate between normal and diseased states: protein microarrays and mass spectrometry. We will concentrate on the latter for the purpose of this chapter.

Mass spectrometry of endogenous human serum peptides using the Ciphergen Biosystems TOF in the MALDI or SELDI mode (Weinberger *et al.*, 2000) as a diagnostic tool and their identification by MALDI-Qq-TOF was successfully demonstrated by Jackowski and coworkers (Takahashi *et al.*, 2001). Later Petricoin and coworkers proposed using only the SELDI pattern of the unidentified peaks as a diagnostic tool (Petricoin *et al.*, 2002a). Biovision (BioVisioN AG, Hannover, Germany) proposed the examination of the MALDI profile of endogenous peptides prepared by reversed phase chromatography against a previously established library of analytes. Their approach is based on identifying patterns of differentially expressed proteins analyzed by computer algorithms, between samples from diseased and nondiseased subjects, without requiring knowledge of the identity of the individual discriminating molecules (Tammen *et al.*, 2003). Since then, many papers have been published on using protein pattern profiling in diagnosing various types of cancer (Table IV) (Adam *et al.*, 2002; Dolios *et al.*, 2003; Ferrari *et al.*, 2000; Koopmann *et al.*, 2004; Kozak *et al.*, 2002; Langridge *et al.*, 1993; Lehrer *et al.*, 2003; Li *et al.*, 2002; Liotta and Petricoin, 2000; Petricoin *et al.*, 2002a,b; Poon *et al.*, 2003; Qu *et al.*, 2002; Rosty *et al.*, 2002; Sasaki *et al.*, 2002; Sauter *et al.*, 2002; Stegner *et al.*, 2004;

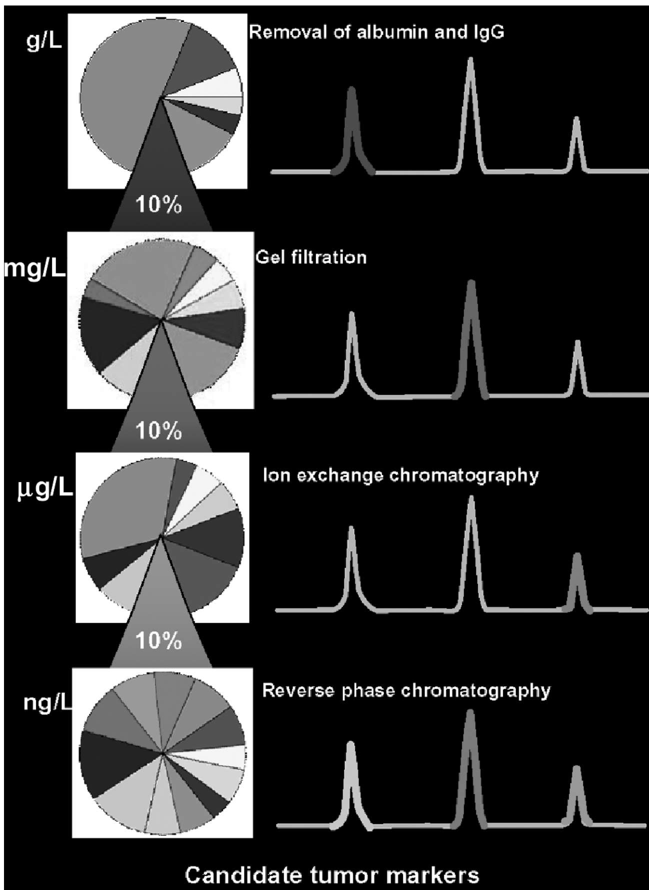

Fig. 3 Schematic representation of successive chromatographic separations for enrichment of fractions with low-abundance proteins. For further discussion see text. (See Color Insert.)

Tammen *et al.*, 2003; Vlahou *et al.*, 2001; von Eggeling *et al.*, 2000; Wadsworth *et al.*, 2004; Wright *et al.*, 1999; Zhukov *et al.*, 2003). The vast majority of the data were generated using SELDI-TOF technology (Ciphergen Biosystems, Fremont, CA). In general, it has been reported that this technology can achieve much higher diagnostic sensitivities and specificities (nearly 100%) compared to classical single biomarkers (Conrads *et al.*, 2004; Powell, 2003). If these findings are reproduced and validated, they could have immediate clinical impact. However, it is important to highlight some limitations of this technique as well and discuss a number of controversial issues surrounding its implementation into clinical practice.

Table IV Protein Pattern Profiling for Cancer Diagnosis

Cancer type	References
Ovarian	Petricoin *et al*., 2002a; Kozak *et al*., 2002
Breast	Liotta and Petricoin, 2000; Li *et al*., 2002; Sauter *et al*., 2002; Stegner *et al*., 2004
Prostate	Adam *et al*., 2002; Lehrer *et al*., 2003; Petricoin *et al*., 2002; Qu *et al*., 2002; Wright *et al*., 1999
Bladder	Langridge *et al*., 1993; Vlahou *et al*., 2001
Pancreatic	Koopmann *et al*., 2004; Rosty *et al*., 2002; Sasaki *et al*., 2002
Head and neck	von Eggeling *et al*., 2000; Wadsworth *et al*., 2004
Lung	Zhukov *et al*., 2003
Colon	Dolios *et al*., 2003
Melanoma	Ferrari *et al*., 2000
Hepatocellular	Poon *et al*., 2003

VI. CURRENT LIMITATIONS OF DIAGNOSTIC MASS SPECTROMETRY

The greatest challenge for proteomic technologies is the inherent complexity of cellular proteomes. Contrary to the genome, the proteome is a dynamic entity, constantly changing in response to cellular or environmental stimuli. Different cells have different proteomes and the proteins within a proteome are structurally quite diverse.

Most of the discussion below will focus on SELDI-TOF mass spectrometry. The possible limitations mentioned here are not unique to this technology, but are relevant to other platforms as well (Diamandis, 2003, 2004a,b; Diamandis and van der Merwe, 2005). The controversy surrounding the method has raised questions as to whether mass spectrometry can meet the standards of reproducibility and performance expected of established clinical tests (Coombes, 2005; Hortin, 2005). A commendable report by Semmes *et al.* (2005) examined the reproducibility between different laboratories and highlighted that this technology does not as yet meet the desired standards to be applied in clinical laboratory practice.

A few common steps are involved in SELDI-TOF procedures. The biological fluid is fractionated with a protein chip, enabling the analysis of subgroups of proteins based on their affinity for a given surface (e.g., normal phase, reverse phase, immobilized metal affinity capture (IMAC), ion-exchange or ligand-binding affinity chromatography) to capture proteins from complex biological samples. After washing, the immobilized proteins are analyzed by SELDI-TOF MS (Fig. 4). The associated shortcomings of the method can be divided into preanalytical, analytical, and postanalytical.

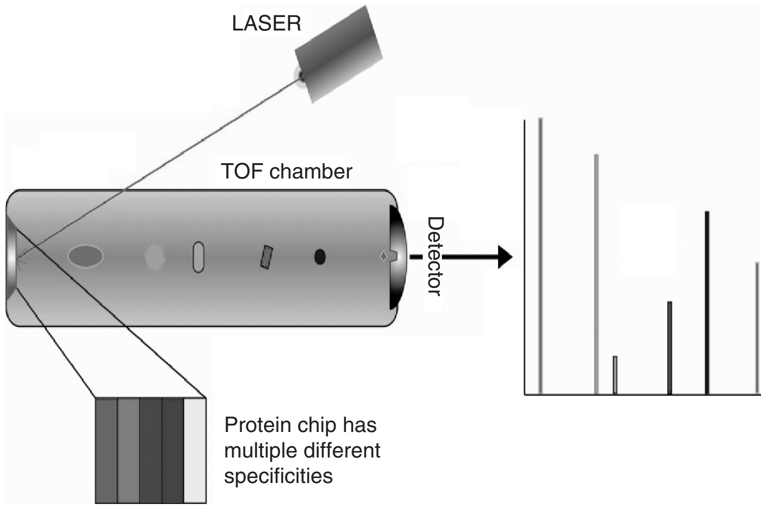

Fig. 4 Schematic representation of SELDI-TOF mass spectrometry. Unfractionated sample is applied to a protein chip, which is coated with various functional groups (1–5) to enable the analysis of a subset of proteins based on affinity for a given surface. Unbound proteins are washed away. A laser beam desorbs and ionizes the proteins, which are cocrystallized with matrix. The mass-to-charge (*m/z*) ratio is determined by the TOF detector and proteomic patterns are analyzed by suitable software. (See Color Insert.)

A. Preanalytical

Many factors influence the concentration of proteins in plasma besides disease. When these are not considered, the detection of medically meaningful changes becomes dubious. The effects of sample storage and processing, sample type (plasma versus serum), patient selection, and different biological variables (gender, age, ethnicity, exercise, menopause, nutrition, drugs, and so on) have as yet not been established yet for mass spectrometric analyses of this type.

B. Analytical

1. DYNAMIC RANGE

The dynamic range of established techniques, such as ELISA, encompasses molecules like albumin in the very high-abundance end (35–50 mg/ml) as well as in the very low-abundance end, for example interleukin-6 (5 pg/ml). The abundance of these two molecules in plasma differs by a factor of 10^{10}. Unbiased protein identification by techniques such as LC/MS/MS have

typical dynamic ranges of only 10^2–10^4, falling short of the requirement for comprehensive proteome mapping by at least 6–8 orders of magnitude. Various fractionation methods (chromatography, immunoaffinity subtraction, preparative isoelectric focusing, or precipitation) improve, but still fall short of the desired dynamic range.

2. ANALYTICAL SENSITIVITY

An important question is whether SELDI-TOF or other associated technologies are sensitive enough to capture putative proteomic changes caused by early stage tumors (Diamandis, 2003, 2004a,b). As currently used, these techniques are unlikely to detect any serum component at concentrations of <1 μg/ml (Diamandis, 2003). This concentration is $\sim$1000-fold higher than the concentrations of known tumor markers in the circulation (see Table III for quantitative comparisons).

3. IDENTIFICATION OF ESTABLISHED CANCER BIOMARKERS

PSA is an established biomarker that can reasonably distinguish cancer from noncancer patients. Free and complexed PSA have molecular masses of $\sim$30 and 100 kDa, respectively, which are well within the current capabilities of mass spectrometry. None of the published studies with breast, prostate, or ovarian cancer identified any of the classical cancer biomarkers as distinguishing molecules. This is likely due to the inadequate sensitivity of currently used protocols, as exemplified in detail elsewhere (Diamandis, 2003).

4. BIAS TOWARD HIGH-ABUNDANCE MOLECULES

Serum contains a wide range of molecules as mentioned earlier (Anderson and Anderson, 2002; Anderson *et al.*, 2004), therefore, competition between high-abundance and low-abundance molecules for immobilization on the protein chip will take place once the sample is applied. For example, PSA concentration in serum of healthy males is in the order of 1 ng/ml compared to a total protein of 80×10^6 ng/ml. When proteins are exposed to the chip, each PSA molecule (or other molecules of similar abundance) will encounter competition for binding to the same matrix by millions of irrelevant molecules of high abundance. Therefore, low-abundance molecules will likely escape binding and detection. Also, the relative amplitudes of peaks in MS spectra will not accurately represent their abundance compared to pure standards. The theoretical sensitivity of MS could be very high (e.g., in the zeptomolar range; Smith and Anderson, 2002), but whether this

is achievable in a complex mixture of high-abundance, as well as low-abundance proteins, remains to be seen. From the available experimental data, current protocols detect primarily or exclusively high-abundance molecules in the concentration range of milligrams per milliliters (Table III).

5. IONIZATION EFFICIENCY

It is not well established whether the same concentration of an informative molecule on the protein chip produces a peak of the same amplitude if it is surrounded by variable amounts of irrelevant proteins that are simultaneously ionized during laser desorption, therefore causing ionization suppression of the relevant molecule. This issue needs to be experimentally examined for each analyte of interest.

6. IDENTITY AND ORIGIN OF DISCRIMINATORY PEAKS

Two different opinions exist in the literature: (1) the identity of the discriminatory peaks produced by MS is not essential and that the diagnostic endpoint is a differentially expressed proteomic profile containing a multitude of molecules reflecting tumor–host interaction, (2) the identity of these peaks is essential, the reason being threefold: to relate their biological connection to cancer, to exclude artifacts originating *ex vivo* during sample handling, and to examine if the findings represent cancer epiphenomena. Most of the discriminatory molecules identified thus far are acute-phase proteins (Table III), released by the liver likely in response to malignancy-associated inflammation (Diamandis, 2003).

7. REPRODUCIBILITY

Some questions and concerns regarding the reproducibility of protein patterns by SELDI-TOF have been raised. There are no systematic studies showing that similar data can be obtained by using different batches of SELDI chips, different technologists, different instrumentation, or at different time points. One hypothesis for the published data is that the differences in serum proteomic patterns between controls and cases are due to the presence of cancer in the latter group. Alternatively, these differences could be due to an unrelated effect, that is, analytical variables or mass spectrometric, bioinformatics, and statistical biases. To date several groups have reported good reproducibility by offline preseparation by C18 partition chromatography prior to MALDI-TOF analysis (Marshall *et al.*, 2003).

Three recent publications dealt with the issue of reproducibility of the serum proteomic test for ovarian cancer (Baggerly *et al.*, 2005; Liotta *et al.*, 2005; Ransohoff, 2005). Baggerly *et al.* (2005) concluded, after analyzing sets of

data produced by Liotta *et al.*, that the discriminatory peaks do not represent biologically important changes in cancer patients and the resulting classification may have arisen by chance. On the other hand, Liotta *et al.* (2005) suggested, as we have proposed previously, to characterize discriminatory peaks so that future classifications are more reproducible and robust. Ransohoff draws attention to biases of experimental designs and suggested that future clinical trials should avoid biological, analytical, statistical, and epidemiological biases (Ransohoff, 2005).

8. ROBUSTNESS

The long-term robustness of this technology needs to be established. Rogers *et al.* (2003) reported that the diagnostic sensitivity in renal cell carcinoma fell from 98 to 41%, tested on two different occasions, 10 months apart. This kind of variability is unacceptable for tests destined to reach the clinic.

C. Postanalytical

1. BIOINFORMATIC ARTIFACTS

SELDI-TOF or associated experiments use a fraction of the clinical samples as a "training set" to derive the interpretation algorithm, while the remaining samples are used as a "test set." Qu *et al.* (2002) pointed out that one of the concerns about learning algorithms is the potential to overfit the data. It is unknown if these algorithms will remain stable over time or when different sets of clinical samples are used. The data of Rogers *et al.* (2003) cast doubt on algorithm stability over time (Poon *et al.*, 2003).

Furthermore, many discriminatory peaks identified to differentiate cancer often have an *m/z* ratio <2000, discarded by many as noise due to matrix effects. Reanalysis of the data of Petricoin *et al.* (2002a) by Sorace and Zhan (2003) and Baggerly *et al.* (2004) revealed that many peaks within the *m/z* range <2000 had powerful discriminatory ability, concluding that a nonbiological bias may best explain the published data and not the presence of cancer.

2. EXTERNAL VALIDATION

The real value of new biomarkers and discovery approaches will ultimately be decided at validation (Ransohoff, 2003, 2004). Efforts to standardize the methodology and test the reproducibility in various laboratories under different clinical settings are underway (Banez *et al.*, 2003; Semmes, 2004;

Table V Some Open Questions Related to Diagnostic SELDI-TOF Technology[a]

1. Identity and serum concentrations of discriminating molecules mostly unknown. These molecules may represent artifacts or cancer epiphenomena
2. Mass spectrometry is a largely qualitative technique
3. Discriminating peaks identified by different investigators for the same disease are different
4. Data are not easily reproducible between laboratories, making validation difficult
5. Optimal sample preparation for the same disease differs between investigators
6. Validated serum cancer markers (PSA, CA125, and so on) that could serve as positive controls are not identified by this technology due to low sensitivity
7. Nonspecific absorption matrices favor extraction of high-abundance proteins and loss of low-abundance proteins
8. Analytical sensitivity of mass spectrometry in a complex mixture (e.g., serum) is unknown
9. No known relationship has been demonstrated between discriminatory molecules and cancer biology

[a]This table was modified from Diamandis (2003) and published with permission from the copyright owners.

Semmes *et al.*, 2005). A summary of the limitations of current MS diagnostic protocols is presented in Table V (Diamandis, 2003).

VII. SUGGESTIONS FOR FUTURE PROGRESS

1. Future investigations should report, whenever possible, the identity of the discriminatory peaks and attempts should be made to link them to cancer biology.
2. Internal controls should be included to correct for peak amplitudes in different experiments.
3. Standardized statistical algorithms that will not vary over time, such as ANOVA, should be used to compare samples and populations (Marshall *et al.*, 2003). Bioinformatic algorithms should be tested periodically to validate their robustness over time.
4. Different bioinformatic algorithms should be compared on the same set of data to determine whether discriminatory peaks and similar diagnostic sensitivity and specificity can be obtained.
5. The analytical sensitivity of mass spectrometry as applied to unfractionated serum samples needs to be determined.
6. Establish whether certain discriminatory peaks originated *ex vivo* (after sample collection) or *in vivo*. Samples from the same patients should be collected with or without proteinase inhibitors and processed in various ways, as described by Marshall *et al.* (2003). In general, serum contains many protein fragments generated during the coagulation cascade.
7. Studies should be performed to establish the effect of pre- and postanalytical variables on proteomic patterns generated, as described earlier. The possibility of bias needs to identified and excluded during each step.

VIII. FUTURE DIRECTION

The complete sequence of the human genome, the development of novel bioinformatic tools, and recent advances in biological mass spectrometry and microarrays sparked optimism that the time has come to discover novel cancer biomarkers. Over the last 3 years, we witnessed an exponential growth of mass spectrometry-based diagnostics with claims of unprecedented clinical sensitivities and specificities. However, various analytical and clinical shortcomings have been recognized. The controversy can be resolved with well-designed validation studies, which are currently underway by investigators, diagnostic companies, and organizations such as the EDRN. Despite these difficulties, it is clear that the opportunities are enormous. For example, mass spectrometry and protein microarrays offer a unique way to simultaneously monitor hundreds to thousands of proteins at the same time. Newer developments may improve the analytical sensitivity of mass spectrometry, allowing measurement of molecules present in biological fluids at very low concentrations. A new discipline called "peptidomics" deals with small peptides in biological fluids which may carry unique information on proteolysis around the cancer microenvironment. Mass spectrometry is ideally suited for high-throughput analysis of a large number of different peptides.

We conclude that mass spectrometry-based diagnostics will continue to grow in the future, with multiparametric analysis of high- and low molecular weight proteins/peptides present in biological fluids in low abundance. Such analysis combined with bioinformatics will eventually lead to novel ways of diagnosing and monitoring cancer. This approach may eventually replace the traditional use of single biomarkers for diagnosis and monitoring of cancer.

REFERENCES

Adam, B.-L., Qu, Y., Davies, J. W., Ward, M. D., Clements, M. A., Cazares, L. H., Semmes, O. J., Schellhammer, P. F., Yassui, Y. M., Feng, Z., and Wright, G. L., Jr. (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.

Aebersold, R., and Goodlett, D. R. (2001). Mass spectrometry in proteomics. *Chem. Rev.* **101,** 269–295.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* **422,** 198–207.

Anderson, N. L., and Anderson, N. G. (2002). The human plasma proteome: History, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1,** 845–867.

Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R. S., Conrads, T. P., Veenstra, T. D., Adkins, J. N., Pounds, J. G., Fagan, R., and Loley, A. (2004). The human plasma proteome: A nonredundant list developed by combination of four separate sources. *Mol. Cell. Proteomics* **3,** 311–326.

Baggerly, K. A., Morris, J. S., and Coombes, K. R. (2004). Reproducibility of SELDI-TOF protein patterns in serum: Comparing data sets from different experiments. *In* "Bioinformatics Advanced Access," pp. 1–9. Oxford, UK.

Baggerly, K. A., Morris, J. S., Edmonson, S. R., and Coombes, K. R. (2005). Signal in noise: Evaluating reported reproducibility of serum proteomic tests for ovarian cancer. *J. Natl. Cancer Inst.* **94,** 307–309.

Banez, L. L., Prasanna, P., Sun, L., Alo, A., Zou, Z., Adam, B. L., McLeod, D. J., Moul, J. W., and Srivastava, S. (2003). Diagnostic potential of serum proteomic patterns in prostate. *J. Urol.* **3**(170), 442–446.

Banks, R. E. (1999). The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis: Preliminary findings. *Electrophoresis* **20,** 689–700.

Bast, R. C., Klug, T. L., St. John, E., Jenison, E., Niloff, J. M., Lazarus, H., Berkowitz, H. S., Leavitt, T., Griffiths, C. T., Parker, L., Zurawski, V. R., and Knapp, R. C. (1983). A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N. Engl. J. Med.* **309,** 883–887.

Bast, R. C., Radvin, P., Hayes, D. F., Bates, S., Fritshce, H., Jessup, J. M., Kemeny, N., Locker, G. Y., Mennel, R. G., and Somerfield, M. R. (2001a). 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: Clinical practice guidelines of the American Society of Clinical Oncology. *J. Clin. Oncol.* **19,** 1865–1878.

Bast, R. C., Radvin, P., Hayes, D. F., Bates, S., Fritshce, H., Jessup, J. M., Kemeny, N., Locker, G. Y., Mennel, R. G., and Somerfield, M. R. (2001b). 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: Clinical practice guidelines of the American Society of Clinical Oncology. *J. Clin. Oncol.* **19,** 4185–4188.

Brenner, S., and Johnson, M. (2000). Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **18,** 630–634.

Caprioli, R. M., Farmer, T. B., and Gile, J. (1997). Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **69,** 4751–4760.

Chaurand, P., Schwartz, S. A., and Caprioli, R. M. (2002). Imaging mass spectrometry: A new tool to investigate the spatial organization of peptides and proteins in mammalian tissue sections. *Curr. Opin. Chem. Biol.* **6,** 676–681.

Cohen, L. S., Escobar, P. F., Scharm, C., Glimco, B., and Fishman, D. A. (2001). Three-dimensional power Doppler ultrasound improves the diagnostic accuracy for ovarian cancer prediction. *Gynecol. Oncol.* **82,** 40–48.

Conrads, T. P., Issaq, H. J., and Veenstra, T. D. (2002). New tools for quantitative phosphoproteome analysis. *Biochem. Biophys. Res. Commun.* **290,** 885–890.

Conrads, T. P., Fusaro, V. A., Ross, S., Johann, D., Vinodh, R., Hitt, B. A., Steinberg, S. M., Kohn, E. C., Fishman, D. A., Whiteley, G., Barrett, J. C., Liotta, L. A., *et al.* (2004). High-resolution serum proteomic features for ovarian cancer detection. *Endocr. Relat. Cancer* **11,** 163–178.

Coombes, K. R. (2005). Analysis of mass spectrometry profiles of the serum proteome. *Clin. Chem.* **51,** 1–2.

Diamandis, E. P. (2003). Point: Proteomic patterns in biological fluids: Do they represent the future of cancer diagnostics? *Clin. Chem.* **49,** 1272–1275.

Diamandis, E. P. (2004a). Analysis of serum proteomic patterns for early cancer diagnosis: Drawing attention to potential problems. *J. Natl. Cancer Inst.* **96,** 353–356.

Diamandis, E. P. (2004b). Mass spectrometry as a diagnostic and a cancer biomarker discovery tool. Opportunities and potential limitations. *Mol. Cell. Proteomics* **3,** 367–378.

Diamandis, E. P., and van der Merwe, D. (2005). Plasma protein profiling by mass spectrometry for cancer diagnosis: Opportunities and limitations. *Clin. Cancer Res.* **11,** 963–965.

Dolios, G., Roboz, J., and Wang, R. (2003). Identification of colon cancer associated protein in plasma using MALDI-TOF mass spectrometry. 51st ASMS Conference on Mass Spectrometry and Allied Topics.

Duffy, M. J., van Dalen, A., Haglund, C., Jansson, L., Klapdor, R., Lamerz, R., Nilsson, O., Sturgeon, C., and Topolcan, O. (2003). Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines. *Eur. J. Cancer* **39,** 718–727.

Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S. R., Weiss, R. A., and Liotta, L. A. (1996). Laser capture microdissection. *Science* **20,** 998–1001.

Etzioni, R., Urban, M., Ramsey, S., McIntosh, M., Schwarz, S., Reid, B., Radich, J., Anderson, G., and Hartwell, L. (2003). The case for early detection. *Nat. Rev. Cancer* **3,** 243–252.

European Group on Tumour Markers (1999). Tumor markers in gastrointestinal cancers-EGTM Recommendations. *Anticancer Res.* **19,** 2785–2820.

Fenn, J. B., Mann, M., Meng, C. K., Eong, S. F., and Whitehouse, C. M. (1989). Electrospray ionization for the mass spectrometry of large biomolecules. *Science* **2456,** 64–71.

Ferrari, L., Seraglia, R., Rossi, C. R., Bertazzo, A., Lise, M., Allegri, G., and Traldi, P. (2000). Protein profiles in sera of patients with malignant cutaneous melanoma. *Rapid Commun. Mass Spectrom.* **14,** 1149–1154.

Fleisher, M., Dnistrian, A. M., Sturgeon, C. M., Lamerz, R., and Wittliff, J. L. (2002). *In* "Tumour Markers: Physiology, Pathobiology, Technology and Clinical Applications" (E. P. Diamandis, H. A. Fritsche, H. Lilja, D. W. Chan, and M. K. Schwartz, Eds.), pp. 33–63. AACC Press, Washington, DC.

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999.

Hampton, G. M., and Frierson, H. F. (2003). Classifying human cancer by analysis of gene expression. *Trend. Mol. Med.* **9,** 5–10.

Hellstrom, I., Raycraft, J., Hayden-Ledbetter, M., Ledbetter, J. A., Schummer, M., McIntosh, M., Drescher, C., Urban, N., and Hellstrom, K. E. (2003). The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res.* **63,** 3695–3700.

Hermeking, H. (2003). Serial analysis of gene expression and cancer. *Curr. Opin. Oncol.* **15,** 44–49.

Hess, J. L. (2003). The Cancer Genome Anatomy Project: Power tools for cancer biologists. *Cancer Invest.* **21,** 325–326.

Hortin, G. L. (2005). Can mass spectrometric protein profiling meet desired standards of clinical laboratory practice. *Clin. Chem.* **51,** 3–5.

Hunt, D. F., and Shabanowitz, J. (1987). Tandem quadrupole Fourier-transform mass spectrometry of oligopeptides and small proteins. *Proc. Natl. Acad. Sci. USA* **84,** 620–623.

Jacobs, I. J., Skates, S. J., MacDonald, N., Menon, U., Rosenthal, A. N., Davies, A. P., Woolas, R., Jeyarajah, A. R., Sibley, K., Lowe, D. G., and Oram, D. H. (1999). Screening for ovarian cancer: A pilot randomized controlled trial. *Lancet* **353,** 1207–1210.

Jellum, E., Bjornson, I., Nesbakken, R., Johansson, E., and Wold, S. (1981). Classification of human cancer cells by means of capillary gas chromatography and pattern recognition analysis. *J. Chromatogr.* **217,** 231–237.

Johnson, M. A., Rohlfs, E. M., and Silverman, L. M. (1994). Proteins. *In* "Tietz Textbook of Clinical Chemistry" (C. A. Burtis and E. R. Ashwood, Eds.), pp. 477–540. WB Saunders Co., Philadelphia.

Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular mass exceeding 10,000 daltons. *Anal. Chem.* **60,** 2299–2301.

Koomen, J. M., Shih, L. N., Coombes, K. R., Li, D., Xiao, L.-C., Fidler, I. J., Abbruzzese, J. L., and Kobayashi, R. (2005). Plasma protein profiling for diagnosis of pancreatic cancer reveals the presence of host response proteins. *Clin. Cancer Res.* **11,** 1110–1118.

Koopmann, J., Zhang, Z., Rosenzweig, J., Fedarko, N., Jagannah, S., Canto, M. I., Yeo, C. J., Chan, D. W., and Goggins, M. (2004). Serum diagnosis of pancreatic adencarcinoma using surface-enhance laser desorption and ionization mass spectrometry. *Clin. Cancer Res.* **10,** 860–868.

Kozak, K. R., Amneus, Z., Rosenzweig, J., Shih, L., Pham, T., Fung, E. T., Sokoll, L. J., and Chan, D. W. (2002). Proteomic approaches to tumor marker discovery. *Arch. Pathol. Lab. Med.* **126,** 1518–1526.

Langridge, J., McClure, T., El-Shakawi, S., Fielding, A. S. K., and Newton, R. (1993). Gas chromatography/mass spectrometric analysis of urinary nucleosides in cancer patients. Potential of modified nucleosides as tumour markers. *Rapid Commun. Mass Spectrom.* **7,** 427–434.

Lehrer, S., Roboz, J., Ding, H., Zhao, S., Diamon, E. J., Holland, J. F., Stone, N. N., Droller, M. J., and Stock, R. G. (2003). Putative protein markers in the sera of men with prostatic neoplasms. *Br. J. U. Intl.* **92,** 223–225.

Li, J., Zhang, Z., Rosenzweig, J., Wang, Y. Y., and Chan, D. W. (2002). Proteomics and bioinformatics approach for identification of serum biomarkers to detect breast cancer. *Clin. Chem.* **48,** 1296–1304.

Lim, M. S., and Elenitoba-Johnson, K. S. J. (2004). Proteomics in pathology research. *Lab. Invest.* **84,** 1227–1244.

Liotta, L., and Petricoin, E. F. (2000). Molecular profiling of human cancer. *Nat. Rev. Genet.* **1,** 48–165.

Liotta, L. A., and Kohn, E. C. (2001). The microenvironment of the tumour-host interface. *Nature* **411,** 375–379.

Liotta, L. A., Ferrari, M., and Petricoin, E. (2003). Clinical proteomics: Written in blood. *Nature* **425,** 905.

Liotta, L. A., Lowenthal, M., Mehta, A., Conrads, T. P., Veenstra, T. D., Fishman, D. A., and Petricoin, E. F. (2005). Importance of communication between producers and consumers of publicly available experimental data. *J. Natl. Cancer Inst.* **97,** 310–314.

Loboda, A. V., and Krutchinsky, A. N. (2000). A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: Design and performance. *Rapid Commun. Mass Spectrom.* **14,** 1047–1057.

Loi, S., Haydon, A. M., Shapiro, J., Schwarz, M. A., and Schneider, H. G. (2004). Towards evidence-based use of serum tumour marker requests: An audit of use in a tertiary hospital. *Int. Med. J.* **34,** 545–550.

Lu, K. H., Patterson, A. P., Wang, L., Marquez, R. T., Atkinson, E. N., Baggerly, K. A., Ramoth, L. R., Rosen, D. G., Liu, J., Hellstrom, I., Smith, D., Harmann, L., *et al.* (2004). Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis. *Clin. Cancer Res.* **10,** 3291–3300.

MacCoss, M. J., Wu, C. C., and Yates, J. R., III (2002). Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal. Chem.* **74,** 5593–5599.

Magklara, A., Scorilas, A., Stephen, C., Kristiansen, G. O., Hauptmann, S., Jung, K., and Diamandis, E. P. (2000). Decreased concentrations of prostate-specific antigen and human glandular kallikrein 2 in malignant versus nonmalignant prostatic tissue. *Urology* **56,** 527–532.

Mann, M., Hendrickson, R. C., and Pandey, A. (2001). Analysis of proteins and proteomics by mass spectrometry. *Ann. Rev. Biochem.* **70,** 437–473.

Marshall, J., and Jankowski, A. (2004). Human serum proteins preseparated by electrophoresis or chromatography followed by tandem mass spectrometry. *J. Proteome Res.* **3,** 364–382.

Marshall, J., Kupchak, P., Zhu, W., Yantha, J., Vrees, T., Furesz, S., Jacks, K., Smith, C., Kireeva, I., Zhang, R., Takahashi, M., Stanton, E., *et al.* (2003). Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J. Proteome Res.* **2,** 361–372.

Menon, U., and Jacobs, I. J. (2000). Recent developments in ovarian cancer screening. *Curr. Opin. Obstet. Gynecol.* **12,** 39–42.

Menon, U., and Jacobs, I. J. (2002). Tumor markers. *In* "Principles and Practice of Gynecologic Oncology" (W. J. Hoskins, C. A. Perez, and R. C. Young, Eds.), 3rd edn. pp. 165–182. Lippincott, Williams & Wilkins, Philadelphia, PA.

Merchant, M., and Weinberger, S. R. (2000). Recent advancements in surface-enhanced laser desertion/ionization-time of flight mass spectrometry. *Electrophoresis* **21,** 1164–1177.

Meyer, T., and Rustin, G. J. (2000). Role of tumour markers in monitoring epithelial ovarian cancer. *Br. J. Cancer* **82,** 1535–1538.

Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., and Roepstorff, P. (2000). Quantitation of peptides and proteins by matrix-assisted laser desorption/ ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun. Mass Spectrom.* **14,** 1226–1232.

Nakanishi, T., Okamoto, N., Tanaka, K., and Shimizu, A. (1994). Laser desorption time-of-flight mass spectrometric analysis of transferring precipitated with antiserum: A unique simple method to identify molecular weight variants. *Biol. Mass Spectrom.* **23,** 220–223.

National Cancer Institute. (2002). Surveillance epidemiology and end results program (online), (cited November 19, 2002), http://seer.cancer.gov/2002.

Oosterhuis, W. P., Bruns, D. E., Watine, J., Sandberg, S., and Horvath, A. R. (2004). Evidence-based guidelines in laboratory medicine: Principles and methods. *Clin. Chem.* **50,** 806–818.

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. *Nature* **405,** 837–846.

Pedrioli, P. G., Eng, J. K., Hubley, R., Vogelzang, M., Deutsch, E. W., Raught, B., Pratt, B., Nilsson, E., Angeletti, R. H., Apweiler, R., Cheing, K., Costello, C. E., *et al.* (2004). A common open representation of mass spectrometry data and its application to proteomics research. *Nat. Biotechnol.* **22,** 1459–1466.

Petricoin, E. F., III, Ardekani, A. M., Hitt, B. A., Levine, P., Fusaro, V. A., Steinberg, S., Mills, G. B., Simcoe, C., Fishman, D. A., Kohn, D. C., and Liotta, L. A. (2002a). Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* **359,** 572–575.

Petricoin, E. F., III, Ornstein, D. K., Paweletz, C. P., Ardekani, A., Hackett, P. S., and Hitt, B. A. (2002b). Serum proteomic patterns for detection if prostate cancer. *J. Natl. Cancer Inst.* **94,** 1576–1578.

Petricoin, E. F., III, Zoon, K. C., Kohn, E. C., Barret, J. C., and Liotta, L. A. (2002c). Clinical proteomics: Translating benchside promise into bedside reality. *Nat. Rev. Drug Disc.* **1,** 683–695.

Polyak, K., and Riggens, G. J. (2001). Gene discovery using the serial analysis of gene expression technique: Implications for cancer research. *J. Clin. Oncol.* **19,** 2948–2958.

Poon, T. C. W., Yip, T., Chan, A. T. C., Yip, C., Yip, V., Mok, T. S. K., Lee, C. C. Y., Leung, T. W. T., Ho, S. K. W., and Johnson, P. J. (2003). Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes. *Clin. Chem.* **49,** 752–760.

Powell, K. (2003). Proteomics delivers on promise of cancer biomarkers. *Nat. Med.* **9,** 980.

Pusch, W., Floccoo, M. T., Leung, S., Thiele, H., and Kostrzewa, M. (2003). Mass spectrometry-based clinical proteomics. *Pharmagenomics* **4,** 463–476.

Qu, Y., Adam, B.-L., Yasui, Y., Ward, M. D., Cazares, L. H., Schellhammer, P. F., Feng, Z., Semmes, O. J., and Wright, G. L., Jr. (2002). Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles disciminated prostate cancer from non-prostate patients. *Clin. Chem.* **48,** 1835–1843.

Ransohoff, D. F. (2003). Cancer. Developing molecular biomarkers for cancer. *Science* **299,** 1679–1680.

Ransohoff, D. F. (2004). Rules of evidence for cancer molecular-marker discovery and validation. *Nat. Rev. Cancer* **4,** 309–314.

Ransohoff, D. F. (2005). Lessons from controversy: Ovarian cancer screening and serum proteomics. *J. Natl. Cancer Inst.* **97,** 315–319.

Reyzer, M. L., Cladwell, R. L., Dugger, T. C., Forbes, J. T., Ritter, C. A., Guix, M., Arteaga, C. L., and Caprioli, R. M. (2004). Early changes in protein expression detected by mass spectrometry predict tumor response to molecular therapeutics. *Cancer Res.* **64,** 9093–9100.

Rogers, M. A., Clarke, P., Noble, J., Munro, N. P., Paul, A., and Selby, P. J. (2003). Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural network analysis: Identification of key issues affecting potential clinical utility. *Cancer Res.* **63,** 6971–6983.

Rosty, C., Christa, L. M., Kuzdzal, S., Baldwin, W. M., Zahurak, M. L., Carnot, F., Chan, D. W., Canto, M., Lillemoe, K. D., Cameron, J. L., Yeo, C. J., Hruban, R. H., *et al.* (2002). Identification of hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein 1 as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology. *Cancer Res.* **62,** 1868–1875.

Sasaki, K., Sato, K., Akiyama, Y., Yanagihara, K., Masaaki, O., and Yamauchi, K. (2002). Peptidomics-based approach reveals the secretion of the 29-residue COOH-terminal fragment of the putative tumor suppressor protein DMBT 1 from pancreatic adenocarcinoma cell lines. *Cancer Res.* **62,** 4894–4898.

Sauter, E. R., Zhu, W., Wassell, R. P., Chervoneva, I., and Du Bois, G. C. (2002). Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. *Br. J. Cancer* **86,** 1440–1443.

Schwartz, S. A., Weil, R. J., Johnson, M. D., Toms, S. A., and Caprioli, M. R. (2004). Protein profiling in brain tumors using mass spectrometry: Feasibility of a new technique for the analysis of protein expression. *Clin. Cancer Res.* **10,** 981–987.

Semmes, O. J. (2004). Defining the role of mass spectrometry in cancer diagnostics. (Editorial). *Cancer Epidemiol. Biomarkers Prev.* **13,** 1555–1557.

Semmes, O. J., Ziding, F., Dam, B.-L., Banez, L. L., Bigbee, W. L., Campos, D., Cazares, L. H., Chan, D. W., Grizzle, W. E., Izbicka, E., Kagan, J., Malik, G., *et al.* (2005). Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. *Clin. Chem.* **51,** 102–112.

Smith, R. D., and Anderson, G. A. (2002). An accurate mass tag strategy for quantitative and high-throughput proteome measurements. *Proteomics* **2,** 513–523.

Sorace, J. M., and Zhan, M. A. (2003). Data review and re-assessment of ovarian cancer serum proteomic profiling. *BMC Bioinformatics* **4,** 24.

Srivastava, S., and Kramer, B. S. (2000). Early detection cancer research network [Editorial]. *Lab. Invest.* **80,** 1147–1148.

Stegner, A. H., Wagner-Mann, C., Du Bois, G. C., and Sauters, E. R. (2004). Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin. Cancer Res.* **10,** 7500–7510.

Sturgeon, C. M. (2001). Tumor markers in the laboratory: Closing the guideline-practice gap. *Clin. Biochem.* **34,** 353–359.

Sturgeon, C. M. (2002). Practice guidelines for tumor marker use in the clinic. *Clin. Chem.* **48,** 1151–1159.

Sullivan, P. M., Etzioni, R., Feng, Z., Potter, J. D., Thompson, M. L., Thornquist, M., Winget, M., and Yasui, Y. (2001). Phases of biomarker development for early detection of cancer. *J. Natl. Cancer Inst.* **93,** 1054–1061.

Takahashi, M., Stanton, E., Moreno, J. I., and Jackowski, G. (2001). Delivering on the promise of proteomics. (cited April 2001), www.synxpharma.com

Tammen, H., Kreipe, H., Hess, R., Kellmann, M., Lehmann, U., Pich, A., Lamping, N., Schulz-Knappe, P., Zucht, H. D., and Lilischkis, R. (2003). Expression profiling of breast cancer cell by differential peptide display. *Breast Cancer Res. Treat.* **79,** 83–93.

Tumor Marker Expert Panel (ASCO). (1996). Clinical practice guidelines for the use of tumor markers in breast and colorectal cancer. *J. Clin. Oncol.* **14,** 2843–2877.

Tuteja, R., and Tuteja, N. (2004). Serial analysis of gene expression (SAGE): Unraveling the bioinformatics tools. *Bioessays* **26,** 916–922.

Tyers, M., and Mann, M. (2003). From genomics to proteomics. *Nature* **422,** 193–197.

Van Dalen, A. (1993). Quality control and standardization of tumor marker tests. *Tumor Biol.* **14,** 131–135.

Villanueva, J., Philip, J., Entenberg, D., Chaparro, C. A., Tanwar, M. K., Holland, E. C., and Tempst, P. (2004). Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal. Chem.* **76,** 1560–1570.

Vlahou, A., Schellhammer, P. F., Mendinos, S., Patel, K., Kondylis, F. I., Gong, L., Nasim, S., and Wright, G. L. J. (2001). Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. *Am. J. Pathol.* **158,** 1491–1502.

von Eggeling, F., Davies, H., Lomas, L., Fiedler, W., Junker, K., Classen, U., and Ernst, G. (2000). Tissue-specific microdissection coupled with protein chip array technologies: Applications in cancer research. *Biotechniques* **29,** 1066–1069.

Von Haller, P. D., Yi, E., Donohoe, S., Vaughn, K., Keller, A., Nesvizhskii, A. I., Eng, J., Li, X. J., Goodlett, D. R., Aebersold, R., and Watts, J. D. (2003a). The application of new software tools to quantitative protein profiling via isotope-coded affinity tag (ICAT) and tandem mass spectrometry: I. Statistically annotated datasets for peptide sequences and proteins identified via the application of ICAT and tandem mass spectrometry to proteins copurifying with T cell lipid rafts. *Mol. Cell. Proteomics* **2,** 426–427.

Von Haller, P. D., Yi, E., Donohoe, S., Vaughn, K., Keller, A., Nesvizhskii, A. I., Eng, J., Li, X. J., Goodlett, D. R., Aebersold, R., and Watts, J. D. (2003b). The application of new software tools to quantitative protein profiling via isotope-coded affinity tag (ICAT) and tandem mass spectrometry: II. Evaluation of tandem mass spectrometry methodologies for large-scale protein analysis, and the application of statistical tools for data analysis and interpretation. *Mol. Cell. Proteomics* **2,** 428–442.

Wadsworth, J. T., Somers, K. D., Stack, B. C. J., Cazares, L., Malik, G., Adam, B. L., Wright, G. L., and Semmes, O. J. (2004). Identification of patients with head and neck cancer using serum protein profiles. *Arch. Otolaryngol. Head Neck Surg.* **130,** 98–104.

Weinberger, S.R, Morris, T. S., and Pawlak, M. (2000). Recent trends in protein biochip technology. *Pharmacogenomics* **1,** 395–416.

Welsh, J. B., Sapinoso, L. M., Su, A. I., Kern, S. G., Wang-Rodriquez, J., Moskaluk, C. A., Frierson, H. F., Jr., and Hampton, G. M. (2001). Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res.* **61,** 5974–5978.

Welsh, J. B., Sapinoso, L. M., Kern, S. G., Brown, D. A., Liu, T., Bauskin, A. R., Ward, R. L., Hawkins, N.J, Quinn, D. I., Russell, P. J., Sutherland, R. L., Breit, S. N., *et al.* (2003). Large scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc. Natl. Acad. Sci. USA* **100,** 3410–3415.

Wilkins, M. R., Gasteiger, E., Wheeler, C. H., Lindskog, I., Sanchez, J. C., Bairoch, A., Appel, R. D., Dunn, M. J., and Hochstrasser, D. F. (1998). Multiple parameter cross-species protein identification using Multi-indent: A world-wide web accessible tool. *Electrophoresis* **19,** 3199–3206.

Winawer, S. J., St. John, D. J., Bond, J. H., Rozen, P., Burt, R. W., Waye, J. D., Kronborg, O., O'Brien, M. J., Bishop, D. T., Kurtz, R. C., *et al.* (1995). Prevention of colorectal cancer: Guidelines based on new data. WHO Collaborating centre for the Prevention of Colorectal. *Cancer Bull. World Health Organ.* **73,** 7–10.

Wright, G. L., Cazares, L. H., Leung, S., Nasim, S., Adam, B., Yip, T., Schellhammer, P. F., Gong, L., and Vlahou, A. (1999). Protein chip surface enhanced laser desorption/ionization

(SELDI) mass spectrometry: A novel biochip technology for detection of prostate cancer biomarkers in complex protein mixtures. *Prost. Cancer Prost. Dis.* **2,** 264–276.

Yanagisawa, K., Shyr, Y., Xu, B. J., Massion, P. P., Larsen, P. H., White, B. C., Roberts, J. R., Edgerton, M., Gonzalez, A., Nadaf, S., Moore, J. H., Caprioli, R. M., *et al.* (2003). Proteomic patterns of tumor subsets of non-small-cell lung carcinoma. *Lancet* **362,** 433–439.

Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001). Proteolytic 18-O labeling for comparative proteomics: Model studies with two serotype of adenovirus. *Anal. Chem.* **73,** 2836–2842.

Yousef, G. M., Polymeris, M. E., Yaccoub, G. M., Scorilas, A., Soosaipillai, A., Popalis, C., Fracchioli, S., Katsaros, D., and Diamandis, E. P. (2003). Parallel overexpression of seven kallikrein genes in ovarian cancer. *Cancer Res.* **63,** 2223–2227.

Zarrinkar, P. P., Mainquist, J. K., Zamora, M., Stern, D., Welsh, J. B., Sapinoso, L. M., Hampton, G. M., and Lockhart, D. J. (2001). Arrays of arrays for high-throughput gene expression profiling. *Genome Res.* **11,** 1256–1261.

Zhang, W., and Chait, B. T. (2000). ProFound: An expert system for protein identification using mass spectrometric peptide mapping information. *Anal. Chem.* **72,** 2482–2489.

Zhang, Z., Bast, R. C., Jr., Yu, Y., Li, J., Sokoll, L. J., Rai, A. J., Rosenzweig, J. M., Cameron, B., Wang, Y. Y., Meng, X. Y., Berchuck, A., van Haaften-Day, C., *et al.* (2004). Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res.* **64,** 5882–5890.

Zhukov, T. A., Johanson, R. A., Cantor, A. B., Clark, R. A., and Tockman, M. S. (2003). Discovery of distinct protein profiles specific for lung tumors and premalignant lung lesions by SELDI mass spectrometry. *Lung Cancer* **40,** 267–279.