

Mass Spectrometry as a Diagnostic and a Cancer Biomarker Discovery Tool

OPPORTUNITIES AND POTENTIAL LIMITATIONS*

Eleftherios P. Diamandis‡

Serum proteomic profiling, by using surfaced-enhanced laser desorption/ionization-time-of-flight mass spectrometry, is one of the most promising new approaches for cancer diagnostics. Exceptional sensitivities and specificities have been reported for some cancer types such as prostate, ovarian, breast, and bladder cancers. These sensitivities/specificities are far superior to those obtained by using classical cancer biomarkers. In this review, I concentrate more on questions that cast doubt on the results reported and propose experiments to investigate these questions in detail, before the technique is used at the clinic. It is clear that the method needs to be externally and thoroughly validated before clinical implementation is warranted. *Molecular & Cellular Proteomics* 3:367–378, 2004.

Our current efforts to combat cancer are not very successful. Despite the recent spectacular advances in molecular medicine, genomics, proteomics, and translational research, mortality rates for the most prevalent cancers have not been significantly reduced. Some of the best available options to combat cancer include primary prevention, earlier diagnosis, and improved therapeutic interventions. We are now witnessing the development of new drugs against cancer that are based on rational instead of empirical designs. There is hope that some of these drugs will prove to be more effective at the clinic than older generations of medicines. In terms of primary prevention, we do not as yet have at hand any robust strategies, because the mechanisms of cancer initiation and progression are still largely unknown.

One of the best strategies to combat cancer now is by early diagnosis and administration of effective treatment (1). Another approach includes close monitoring of the cancer patient after initial treatment (usually surgery) to detect early relapse and then prescribe additional therapy. A third valuable approach would be the stratification of patients into subgroups that respond better to different types of treatment

(individualized therapy). Medical imaging and serum or tissue biomarkers are valuable tools for monitoring these patients in order to optimize clinical outcomes.

In this review, I will concentrate on mass spectrometry as a diagnostic and cancer biomarker discovery tool. Much has been published on this technology, and excellent reviews have already been prepared (2–12). My presentation will be biased toward underlining potential limitations that have not been adequately addressed in the already existing extensive literature.

MASS SPECTROMETRY

Mass spectrometry has been used as a diagnostic tool in clinical laboratories for many decades. This technology has been coupled with gas chromatography (GC/MS)¹ and has been used with success for the identification and quantification of relatively small molecules (with molecular mass <1,000 Da). Such molecules could be highly informative in newborn screening programs (13), toxicological and forensic applications (14), for delineating various types of inborn errors of metabolism (15), for detecting doping of athletes (16), etc. Over the last 15 years, we have seen a resurgence of this technology for studying larger molecules such as nucleic acids and proteins. These new applications became possible mainly due to the development of novel methodologies to effectively volatilize and ionize proteins and nucleic acids, by using various chemicals (matrices) and lasers (e.g. matrix-assisted laser desorption/ionization, MALDI) or electrospray ionization (ESI). The ability to measure with high accuracy mass-to-charge ratio, providing spectra of very high resolution, and the development of tandem mass spectrometry (MS/MS) to obtain *de novo* protein sequence information has further enhanced the applications of this technology in proteomics. Coupling of mass spectrometers to liquid chromatography (LC/MS) further expanded the discriminatory power of the method. Mass spectrometry is now one of the most powerful proteomic tools (17). Even more spectacular advances in mass spectrometry should be expected, with fur-

From the Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada and Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada.

Received, February 27, 2004

Published, MCP Papers in Press, February 28, 2004, DOI 10.1074/mcp.R400007-MCP200

¹ The abbreviations used are: GC/MS, gas chromatography/mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; ESI, electrospray ionization; MS/MS, tandem mass spectrometry; LC/MS, liquid chromatography/mass spectrometry; SELDI-TOF, surface-enhanced laser desorption/ionization time-of-flight; PSA, prostate-specific antigen; ELISA, enzyme-linked immunosorbent assay; PSMA, prostate-specific membrane antigen.

TABLE I
Some established cancer biomarkers

Biomarker ^a	Cancer type ^b
α-Fetoprotein (AFP)	Hepatoma; testicular cancer
Carcinoembryonic antigen (CEA)	Colon; breast; lung; pancreatic
PSA	Prostate
CA125	Ovarian
CA15.3	Breast
CA19.9	Gastrointestinal
Immunoglobulins	B cell dyscrasias
Choriogonadotropin (hCG)	Testicular cancer; trophoblastic tumors
Steroid hormone receptors	Breast

^a All of these markers are used as aids in diagnosis, prognosis, and monitoring of therapy; steroid hormone receptors are used for predicting therapeutic response to antiestrogens.

^b All markers measured in serum except steroid hormone receptors, which are measured in cancer tissues.

ther improvements in resolution and detectability. With this in mind, it is not surprising that many scientists have decided to use mass spectrometry either as a diagnostic tool or as a cancer or other disease biomarker discovery platform (2–12).

I will need to emphasize at this point that the critical discussion to follow is not directed against either mass spectrometry or to the field of proteomics in general. In fact, these methods and fields of investigation, used appropriately, may indeed succeed in discovering new diagnostic modalities for cancer and other diseases, as well as contribute to the better understanding of the pathogenesis of such diseases. The Human Proteome Organization (HUPO, www.hupo.org) is focusing on the identification of large numbers of proteins in complex mixtures, including serum and other biological fluids (17). It is expected that these efforts will finally lead to the identification of new potential biomarkers for cancer and other diseases. HUPO also intends to standardize the methodology so that the results obtained with these techniques are robust and reproducible among laboratories.

Most of the discussion below will focus on one proteomic platform used extensively in diagnostics, known as surface-enhanced laser desorption/ionization-time-of-flight (SELDI-TOF) mass spectrometry. This technique is based on the pretreatment of a biological fluid or tissue extract with various proteomic chips, performing protein extractions based on hydrophobic, ion-exchange, metal binding, or other interactions. The bound proteins are then subjected to mass spectrometric analysis. The derived information can be used for either diagnosis or for identifying potential biomarkers that could then be further validated with alternative technologies. These issues will be discussed in detail below.

CANCER BIOMARKERS

A handful of cancer biomarkers are currently used routinely for population screening, disease diagnosis, prognosis, monitoring of therapy, and prediction of therapeutic response. Some established biomarkers are listed in Table I. Although it is highly desirable to have biomarkers suitable for population screening and early diagnosis, none of the biomarkers listed

in Table I has adequate sensitivity, specificity, and predictive value for population screening. Even prostate-specific antigen (PSA), which has been approved for population screening by the Food and Drug Administration, is not universally accepted for this application. The reasons for biomarker failure in population screening settings are multiple and fall outside the scope of this review. It will suffice to mention that poor specificity leads to many false-positive results. In population screening, disease prevalence is another important parameter; diseases of low prevalence (like ovarian cancer) will require outstanding diagnostic test specificity (>99%) for the test to be considered viable (18). It can be concluded that none of the individual biomarkers currently at hand can fulfill the requirements of population screening for cancer. Biomarkers are clinically recommended mainly for monitoring the effectiveness of therapeutic interventions. Some biomarkers are also invaluable tools for early diagnosis of cancer relapse, which may trigger additional treatments before the appearance of clinical symptoms.

With current cancer biomarkers, much is left to be desired in terms of clinical applicability. We need new cancer biomarkers that will further enhance our ability to diagnose, prognose, and predict therapeutic response in many types of cancer. Because biomarkers can be analyzed relatively non-invasively and economically, it is worth investing in discovering more biomarkers in the future. The completion of the Human Genome Project has raised expectations that the knowledge of all genes and proteins will lead to the identification of many candidate biomarkers for cancer and other diseases. This prediction still needs to be realized. Among specialists in the field, the prevailing view is that the most powerful single cancer biomarkers may have already been discovered (e.g. those shown in Table I). Likely, we are now bound to discover biomarkers that are less sensitive or specific but that could be used in panels, in combination with powerful bioinformatic tools (such as artificial neural networks, logistic regression, etc.), to devise diagnostic algorithms with improved sensitivity and specificity (19, 20). These efforts are currently ongoing.

GENERAL STRATEGIES FOR DISCOVERING NEW
CANCER BIOMARKERS

Most of the currently used cancer biomarkers were discovered following development of novel analytical techniques, such as immunological assays and the monoclonal antibody technology. It was then found that these molecules were elevated in biological fluids from cancer patients in comparison to normal subjects. Many cancer biomarkers were discovered by immunizing animals with extracts from tumors or cancer cell lines, and then screening for monoclonal antibodies that recognize “cancer-associated” antigens. More recently, and with the completion of the Human Genome Project, many researchers hypothesized that the best cancer biomarkers will likely be secreted proteins (21); about 20–25% of all cell proteins are secreted. However, this is not an absolute requirement because a number of classical cancer biomarkers (e.g. CEA, Her2-neu) are cell membrane-bound, but their extracellular domains are shed into the circulation. Other groups, including our own, are using bioinformatics, such as digital differential display and *in silico* Northern blotting, to compare gene expression between normal and cancerous tissues to identify overexpressed genes (22). Although one of the prevailing hypotheses in new biomarker discovery is that the most promising biomarkers should be overexpressed proteins, this is not generally true for some of the best known cancer biomarkers such as PSA (23). Overexpressed genes are now identified experimentally by using microarrays. Some of these genes have been proposed as candidate cancer biomarkers (24–26). Despite this reasonable hypothesis, very few cancer biomarkers have been discovered by using this approach (26, 27). We followed another approach, in which we postulated that if a molecule is already a known-cancer biomarker, members of the same family of genes/proteins may also constitute novel biomarkers. We have since shown that kallikreins, a group of serine proteases with high homology at both the DNA and protein levels (this family includes PSA), are candidate biomarkers for ovarian, prostate, and breast cancers (28, 29).

Over many years of developing cancer biomarkers, we came to understand that a molecule may become a practical serological biomarker if it has certain characteristics, *i.e.* it is a secreted or shed protein and has the ability to diffuse into the circulation during tumor development and progression, through either angiogenesis or invasion of surrounding tissues and vasculature by cancer cells. Preferably, such proteins should be stable (not degraded) and not bound to inhibitors that could interfere with their measurement. The experience with the classical biomarkers has taught us many lessons on the dynamic relationships between the patient and biological phenomena related to biomarkers such as appearance in the circulation, cleavage, binding to serum proteins, degradation, modification, elimination half-life, etc. In this review, I will use PSA as an example to compare what we know from such molecules with mass spectrometric approaches for diagnostics.

MASS SPECTROMETRY AS A CANCER BIOMARKER DISCOVERY
AND DIAGNOSTIC TOOL

Petricoin *et al.* have pioneered the use of mass spectrometry as a diagnostic tool (30). They suggested that this approach represents a paradigm shift in cancer diagnostics, based on complex mass spectrometric differences between proteomic patterns in serum between patients with or without cancer identified by bioinformatics. Their premise is that no matter what the nature of these molecules are, their potential to discriminate between these two conditions should be further exploited. The central hypothesis of this approach is as follows: protein or protein fragments produced by cancer cells or their microenvironment may eventually enter the general circulation. Then, the concentration (abundance) of these proteins/fragments could be analyzed by mass spectrometry and used for diagnostic purposes, in combination with a mathematical algorithm (30).

The vast majority of the currently available data have been produced by using the SELDI-TOF technology, marketed by Ciphergen Biosystems (Fremont, CA). Ciphergen claims that over 200 papers have already been published with this technology. The types of cancers that have been examined include ovarian, prostate, breast, bladder, renal, and others, and the biological fluids analyzed include serum, urine, cerebrospinal fluid, nipple aspirate fluid, etc. The apparent successes with this technology have been recently reviewed by many investigators (2–12). In general, it has been suggested that this technology can achieve much higher diagnostic sensitivity and specificity (approaching 100%) in comparison to the classical cancer biomarkers (31). The technology’s potential has been expanded to other diseases such as Alzheimer’s disease, Creutzfeldt-Jakob disease, renal allograft rejection, etc. (32–34).

The analytical procedure with this technology involves a few common steps. The biological fluid of interest is first interacted with a protein chip that incorporates some kind of an affinity separation between “noninformative” and “informative” proteins. After washing, the immobilized (and fortunately mostly informative) proteins can be studied by using SELDI-TOF mass spectrometry. Two types of data have been reported in the literature: 1) discriminating peaks of unknown identity that are different in amplitude (increased or decreased) between normal individuals and patients with cancer; and 2) data in which at least some of these peaks have been positively identified (see below). Computer algorithms have been used to analyze these multidimensional data to demonstrate that a pattern consisting of several peaks (from tens to thousands) is sufficiently different between the two groups of subjects. In this review, I will not comment much on peaks that have not been positively identified, because nothing is known about them, except that their heights go up or down in the disease state. I will use the few positively identified molecules to draw comparisons between them and the classical cancer biomarkers.

The extraordinary data presented in the literature with this new approach were welcomed by scientists, the press, the public, and even by politicians (31, 35). This technology is now seen as the most promising way of diagnosing early cancer (35). Clinical trials are now underway and will reveal, in a blinded fashion, if these data can be reproduced and if they are robust enough for clinical use. In the following paragraphs, I will concentrate on issues that have not been adequately addressed and raise concerns that at least some of this data may not be accurate or expected on theoretical grounds.

The use of SELDI-TOF technology as a cancer biomarker discovery tool (as opposed to a cancer diagnostic tool) is straightforward. The discriminatory peaks, if positively identified, may represent molecules that could be measured with simpler and cheaper techniques for the purpose of diagnosing cancer. For example, some investigators postulate that such molecules may be routinely quantified by using enzyme-linked immunosorbent assay (ELISA) technologies. In practice, very few, if any, of the SELDI-TOF identified novel candidate biomarkers have been validated by using alternative technologies.

POTENTIAL LIMITATIONS

Liotta *et al.* hypothesized that the relative cellular abundance of tens of thousands of different proteins, along with their cleaved or modified forms, is a reflection of ongoing physiological and pathological events. They further postulate that as tissues are perfused by blood and lymph, proteins and protein fragments, passively or actively, enter the circulation. Thus, the complex chemistry of the tumor-host microenvironment should generate unique signatures in the blood microenvironment. I agree with this statement. The major question here is if these putative proteomic changes in the blood can be captured by the SELDI-TOF technology, as applied in the published papers. In my opinion, it is highly unlikely that a small and localized tumor will be able to modify the serum proteomic pattern to a degree that can be recognized by the SELDI-TOF technique. As I will further elaborate later, SELDI-TOF, and other proteomic technologies based on mass spectrometry, may not be sensitive enough to detect the low-abundance “signature” molecules that are released by a few tumor cells or their microenvironment into the circulation. I do believe that informative molecules originating from tumor cells or their microenvironment may indeed be present in biological fluids and that their identification may lead to the discovery of potential new biomarkers.

The identification of these molecules will likely require ultrasensitive techniques capable of measuring concentration ranges 10^{-12} mol/liter or lower (far lower than those achieved by current SELDI-TOF protocols, see below).

An alternative hypothesis for the observed differences in proteomic patterns in serum between normal individuals and cancer patients may be the detection of high-abundance molecules that are not produced by the tumor cells but rather represent epiphenomena of tumor presence. For example, it

has been postulated by this author that at least some of the detected molecules represent acute-phase reactants that are released into the circulation by the liver and other organs (36, 37). It has been shown as early as 30–40 years ago that such molecules are not specific for the presence of any cancer, and for this reason they have not been used in clinical practice for cancer diagnosis or monitoring, although their concentrations may be elevated in serum of some cancer patients (38).

DISCREPANCIES BETWEEN RESEARCH GROUPS

As it currently stands, SELDI-TOF technology requires pretreatment of a small amount of serum with SELDI protein chips. These protein chips have either 8 or 16 spots containing a specific chromatographic surface. Currently available surfaces are based on either hydrophobic, ion-exchange, metal affinity, or normal phase chromatography. It is also possible, but not widely utilized at the moment, to immobilize more specific reagents such as antibodies, receptors, DNA, etc. For diagnostics, one would expect that the discriminatory peaks for one cancer type can be identified by using preferentially one of these surfaces. After surveying the literature for prostate and ovarian cancer diagnostics, I identified five papers that used SELDI-TOF technology for prostate cancer (39–43) and two papers for ovarian cancer (30, 44). The different groups found that different proteomic chips may be optimal for disease diagnosis. Metal affinity (IMAC-Cu), hydrophobic (C16 or H4), or weak cation exchange (WCX2) chips were used for prostate cancer. In two of the studies (40, 41), the same mass spectroscopic data were used by the same group but different bioinformatic tools were employed to analyze them. A summary of the prostate and ovarian cancer studies are presented in Tables II and III. The following points are relevant. The distinguishing peaks between cancer and non-cancer patients are very different between the various groups. In fact, none of the distinguishing peaks between the four different research groups for prostate cancer agree with each other. The only agreements were two peaks for distinguishing non-cancer *versus* cancer from the same group of investigators and the same datasets (40, 41). A different bioinformatic analysis revealed other discriminatory peaks between the two studies from the same group (41). Similar discrepancies are seen with ovarian cancer (Table III). How could these discrepancies be explained? One hypothesis is that serum may indeed contain a huge number of discriminatory molecules between cancer and non-cancer patients and that the chance of two groups finding the same discriminatory peaks is very small. Another explanation may be methodological differences in which different chips were used to immobilize the candidate discriminatory peaks. This is not a likely hypothesis because Banez *et al.*, Adam *et al.*, and Qu *et al.* (40–42) used the same protein chip (IMAC-Cu), yet they came up with different discriminatory peaks (Table II). In my opinion, it will be highly unlikely that a small, localized tumor and its microenvironment will generate such diverse populations of

TABLE II
Comparison of five reports for prostate cancer diagnosis based on SELDI-TOF technology^a

Study	Chip type	Distinguishing peaks m/z^b	Diagnostic sensitivity and specificity
Petricoin <i>et al.</i> (39) Adam <i>et al.</i> (40)	Hydrophobic C16 IMAC-Cu	2,092, 2,367, 2,582, 3,080, 4,819, 5,439, 18,220 4,475, 5,074, 5,382, 7,024 , <u>7,820</u> , 8,141, 9,149, 9,507, 9,656	95%; 78–83% 83%; 97%
Qu <i>et al.</i> (41)	IMAC-Cu	Non-cancer vs cancer: 3,963, 4,080, 6,542, 6,797, 6,949, 6,991, 7,024 , 7,885, 8,067, 8,356, 9,656 , 9,720 Healthy vs BPH ^c : 3,486, 4,071, 4,580, 5,298, 6,099, 7,054, <u>7,820</u> , 7,844, 8,943	97–100%; 97–100%
Banez <i>et al.</i> (42)	WCX2 IMAC-Cu	3,972, 8,226, 13,952, 16,087, 25,167, 33,270 3,960, 4,469, 9,713, 10,266, 22,832	63%; 77% 66%; 38%
Lehrer <i>et al.</i> (43)	Hydrophobic H4	Cancer and BPH vs controls: 15,200, 15,900, 17,500 Cancer vs BPH 15,200 15,900 17,500	100% (specificity) 82%; 67% 82%; 100% 64%; 67%

^a This table is modified and expanded from Refs. 46 and 68.

^b m/z ratios were rounded to whole numbers for simplicity. m/z ratios in bold represent those identified by Adam *et al.* (40) and Qu *et al.* (41) for differentiating cancer from non-cancer patients. The underlined m/z ratio represents a peak identified by Adam *et al.* (40) for differentiating cancer from non-cancer patients and by Qu *et al.* (41) for differentiating healthy individuals from patients with benign prostatic hyperplasia.

^c BPH, benign prostatic hyperplasia.

TABLE III
Comparison of two reports for ovarian cancer diagnosis based on SELDI-TOF technology

Study	Chip type	Distinguishing peaks, m/z^a	Diagnostic sensitivity and specificity ^b
Petricoin <i>et al.</i> (30) Kozak <i>et al.</i> (44)	C-16 SAX2 ^c	534, 989, 2,111, 2,251, 2,465 Screening biomarker panel ^d : 4,400, 15,900, 18,900, 23,000, 30,100 Validation biomarker panel 1 ^e : 3,100, 13,900, 21,000, 79,000, 106,700 Validation biomarker panel 2 ^f : 5,100, 16,900, 28,000, 93,000	100%; 95% 96%; 83% 82%; 95% 73%; 95%

^a m/z ratios were rounded to whole numbers for simplicity.

^b Somewhat different values were obtained for different sets of samples; see original papers for more details.

^c Strong anion-exchange chip.

^d Differentiates healthy from neoplastic (benign and malignant) disease. In this paper, the molecular masses of the markers are given instead of m/z ratios. All values were reported with one decimal point accuracy.

^e Differentiates healthy and benign disease from malignant disease.

^f In Ref. 30, there is a limit to the molecular mass or m/z ratio monitored (up to 20,000).

informative peptides/proteins in the circulation. Another important difference, displayed in Table II, refers to the diagnostic sensitivity and specificity. Contrast to Qu *et al.*, Petricoin *et al.*, and Adam *et al.*, Banez *et al.* reported that at least with the IMAC-Cu proteomic chips, their sensitivity and specificity was only 66 and 38%, significantly inferior than the other three studies.

What needs to be done to investigate these discrepancies further? First, the experiments should be independently repeated by other laboratories. Second, these validation studies should be done with the older (Ciphergen) and with higher resolution instruments, various batches of proteomic chips, and by using different bioinformatic tools. Also, internal controls (such as already validated classical discriminatory can-

cer biomarkers) should be incorporated to validate the actual analytical sensitivity of the technology (see below).

FAILURE TO IDENTIFY ESTABLISHED CANCER BIOMARKERS

We currently have at hand validated cancer biomarkers that can reasonably distinguish between cancer and non-cancer patients. For example, PSA can be used as a biomarker for a group of patients without cancer (PSA < 1 $\mu\text{g/liter}$) and patients with histologically confirmed prostate cancer and PSA > 10 $\mu\text{g/liter}$. Because free PSA and complexed PSA have molecular masses of ~30 kDa and 100 kDa, respectively, these masses are well within the current capabilities of mass spectrometers (43). Validation of this technology will be highly enhanced if it is shown that one of the discriminatory

TABLE IV

Concentration of some abundant proteins, putative new cancer biomarkers identified by SELDI-TOF, and classical cancer biomarkers in serum^a

Compound	Approximate concentration [pmol/liter]	Biomarker for cancer type	Reference
Albumin	600,000,000	–	38
Immunoglobulins	30,000,000	–	38
C-reactive protein	40,000	–	38
SELDI-TOF proteins ^b			
Apolipoprotein A1	40,000,000	Ovarian	49
Transthyretin (prealbumin) fragment	6,000,000	Ovarian	49
Inter- α -trypsin inhibitor fragment	4,000,000	Ovarian	49
Haptoglobin- α -subunit	1,000,000	Ovarian	50
Vitamin D-binding protein	10,000,000	Prostate	51
Serum amyloid A protein	20,000,000	Nasopharyngeal	52
Classical tumor markers			
Alpha-fetoprotein	150	Hepatoma, testicular	38
PSA	140	Prostate	38
Carcinoembryonic antigen	30	Colon, lung, breast, pancreatic	38
Human choriongonadotropin	20	Testicular, choriocarcinoma	38
Human choriongonadotropin- β subunit	2	Testicular, choriocarcinoma	38

^a This table is modified and expanded from Ref. 68.

^b Apolipoprotein A1 is produced in liver and intestine; all other proteins are mainly produced in the liver (38).

peaks identified in prostate cancer is PSA and its subfractions. The same comment applies for other validated cancer biomarkers. Surprisingly, in none of the published studies with breast, prostate, or ovarian carcinoma have the classical cancer biomarkers been identified. I believe that the inability to identify these classical cancer biomarkers is due to the low sensitivity of the SELDI-TOF approach. Until validated serum internal controls are used with this technology, the results obtained, and the sensitivity of the method, should remain in question.

BIAS OF SELDI-TOF TECHNOLOGY TOWARD HIGH-ABUNDANCE MOLECULES

The current method of performing SELDI-TOF experiments with unfractionated serum includes exposure of serum to the protein chip, washing, and then identification of the immobilized molecules by using MALDI-TOF instrumentation. The solid phases currently in use (mentioned earlier) are not specific for any type of protein. Because serum contains a tremendous array of extremely high-abundance (e.g. albumin) and very-low-abundance molecules (range of concentrations vary by a factor of 10^6 - to 10^9 -fold) (45), it will be highly unlikely that the most informative, low-abundance molecules will be able to immobilize on such chips. Simply, they will likely be competed out by high-abundance, noninformative molecules. For example, in serum, the PSA concentration in healthy males is $\sim 1 \mu\text{g/liter}$, whereas the total protein concentration is in the order of $80,000,000 \mu\text{g/liter}$. 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. It would thus seem very unlikely that molecules with very low abundance will ever be detected by this method. The exper-

iments to prove or disprove these proposals have been previously outlined by this author in a separate editorial, but to my knowledge they have not as yet been reported (46).

A previous report by Wright *et al.* claimed that four classic prostatic biomarkers, including free and complexed PSA, could be detected by mass spectrometry in various biological fluids and tissue extracts, including seminal plasma, prostatic extracts, and serum (47). However, the masses assigned to free or complexed PSA may have originated from other molecules with a similar molecular mass. Furthermore, the presence of other molecules, such as salts, could cause a mass shift, thus complicating the interpretation further. These authors, in their efforts to show a quantitative relationship between peak area and PSA concentrations, constructed linear calibration curves but at PSA concentrations between 1,000 and $50,000 \mu\text{g/liter}$. Such concentrations are rarely or never seen in clinical practice, even in sera from patients with highly metastatic prostate cancer. On the same point, other authors reported prostate-specific membrane antigen (PSMA) concentrations in serum (this is another prostatic-specific molecule) by using a SELDI-TOF approach in combination with an immobilized antibody. The reported concentrations of PSMA in serum (100 – $500 \mu\text{g/liter}$; ~ 500 times higher than the secreted protein PSA) are surprisingly high and need to be validated by ELISA-type methodologies, given that this molecule is a membrane-bound protein (48).

I have compiled a list of SELDI-TOF-identified molecules in serum that are thought to be discriminatory between normal stage and cancer (Table IV). Clearly, these candidate serum biomarkers are very-high-abundance molecules known to be produced mainly by the liver.

For example, Zhang *et al.* (49) identified three discriminatory

peaks in ovarian cancer: apolipoprotein A1, transthyretin (pre-albumin) fragment, and inter- α -trypsin inhibitor. Ye *et al.* discovered haptoglobin- α subunit for ovarian cancer (50), and Hlavaty *et al.* discovered vitamin D-binding protein for prostate cancer (51). More recently, Cho *et al.* identified serum amyloid A protein for nasopharyngeal carcinoma (52). Table IV presents the comparative serum concentrations of these putative tumor markers and of classical tumor markers, such as α -fetoprotein and PSA.

A number of the “new” tumor biomarkers discovered by SELDI-TOF technology were, in fact, originally identified more than 30 years ago by classical techniques (*e.g.* haptoglobin- α subunit for ovarian cancer) (53) but were deemed useless for clinical diagnosis because of their low sensitivity and specificity (54, 55). Just to illustrate this point further, I performed a MEDLINE search using the keywords “haptoglobin” and “cancer” and identified 571 papers published from 1965 to 2003. Haptoglobin was reported since 1966 to be elevated in the following malignancies: leukemias, Hodgkin’s disease, Burkitt’s lymphoma, multiple myeloma, neuroblastoma, melanoma, glioma, and cancers of the cervix, genitals, stomach, breast, liver, kidney, ovaries, lung, endometrium, colon, prostate, gallbladder, bladder, head and neck, brain, and larynx. The same comments applies to serum amyloid A protein (52). It is clear that haptoglobin- α subunit or other acute-phase reactants are not specific cancer biomarkers.

IDENTITY AND ORIGIN OF DISCRIMINATORY PEAKS

Immediately after publication of the first report of SELDI-TOF-based diagnostics for ovarian cancer (30), I urged the authors to positively identify the discriminatory peaks so that their serum elevation or decrease in cancer is better understood (36). Efforts to identify these discriminatory peaks have been minimal. Liotta *et al.* suggested that knowledge of peak identity should not be essential and that this technology represents a new diagnostic paradigm (56). To date, the identity of the five “discriminatory” peaks for ovarian cancer remains elusive (30). Fortunately, HUPO has currently identified as one of their goals to characterize the serum proteome. Also, newer instrumentation is now capable of identifying the discriminatory peaks by using tandem mass spectrometry. As more peaks are positively identified, we will be able to further understand what this technology really detects and if indeed the identified molecules can be confirmed independently to be potential cancer diagnostic markers by using other methodologies (*e.g.* ELISA). As mentioned, currently identified molecules by this technology are of very high abundance, are mostly produced by the liver, and many are acute-phase reactants (Table IV).

TECHNICAL CAVEATS AND METHODOLOGICAL DETAILS

It is important to understand how this method is used in order to identify possible deficiencies. Routinely, 1–3 μ l of serum, either diluted or undiluted, is added to the activated

surface of the protein chip and incubated. The chip is then washed, air-dried, and treated with an ultraviolet-absorbing agent (*e.g.* sinapinic acid, also known as “matrix”) and then dried. The chip is then inserted into the mass spectrometer for SELDI-TOF mass spectrometric analysis. A critical question here is if the protein chip has the capacity to bind quantitatively all proteins present in the sample. Clearly, the answer is no. Then, what binds to the chip will depend on the total protein of the sample, the abundance of various competing proteins for the solid phase, and the properties of the chip (such as hydrophobicity, ion-exchange capacity, metal binding, etc.). Without knowing the abundance of competing proteins, and given the limited capacity of the matrix, what is finally retained on the surface may be quite variable between different clinical samples. It is thus likely that an informative molecule of the same abundance in two clinical samples may be detected at different abundances simply due to the presence of different amounts of noninformative competing molecules. Thus, the relative amplitudes of peaks in mass spectrometric spectra should be considered as “semi-quantitative” at best. Also, as mentioned earlier, the competitive nature of the binding will likely exclude low-abundance molecules due to preferential binding of high-abundance molecules with similar physicochemical properties.

These issues have not been adequately addressed in any of the published papers. Useful experiments could include spiking of known molecules in serum that is devoid of them. For example, spiking female serum with PSA or other molecules that have been tagged with stable isotopes may help to answer these questions. Similar experiments could delineate the detection limit of the methodology as it applies to SELDI-TOF procedures. Spiking with synthetic peptides would be another option.

Ideally, this method could work quantitatively if the surfaces used for molecule immobilization are either specific for certain proteins (*e.g.* antibodies or other binders) or they have enough capacity to quantitatively bind all the proteins applied to the sample.

It is also important to address the issue of ionization efficiency. Would the same concentration of an informative molecule on the protein chip produce a peak of the same amplitude if it is surrounded by variable amounts of irrelevant proteins that are also ionized during laser desorption? One would expect that the ionization will likely be affected by the presence of other molecules in the mixture, further contributing to the qualitative nature of the measurement.

As further stressed by Aebersold and Mann (17) in both MALDI and ESI-MS ionization, the relationship between the amount of analyte present and measured signal intensity is complex and incompletely understood. Mass spectrometers are therefore inherently poor quantitative devices. Furthermore, the ion current of a peptide is dependent on a multitude of variables that are difficult to control, and this measure is not a good indication of peptide abundance. To conclude, mass spectrometric analysis is not quantitative at present.

Regarding sensitivity of mass spectrometry, this is difficult to define because sensitivity is heavily dependent on the machine used (and these are rapidly changing over time) as well as the actual procedures performed before sample introduction into the instrument. For example, some procedures include extensive purification and preconcentration of samples, while in others (including SELDI-TOF analysis) the samples are minimally treated. Nevertheless, in a recent study of global analysis of protein expression in yeast, Ghaemmaghami *et al.* (57) compared mass spectrometry and protein tagging methodologies combined with either Western blotting or green fluorescent protein microscopy. They found that tagging technologies and fluorescence microscopy were able to detect a total of 4,517 proteins with more than 90% overlap. In contrast, a recent study using mass spectrometry and isotope labeling succeeded in quantitatively monitoring changes in the abundance of only 688 yeast proteins (58). The authors concluded that mass spectrometry is capable of detecting abundances of proteins with >50,000 molecules per cell but it was not sensitive in detecting proteins with abundances of <5,000 molecules per cell. Thus, the current inherent relative insensitivity of mass spectrometry, in comparison to Western blotting and fluorescence microscopy, combined with the fact that low-abundance proteins may not bind to the biochip, will make the detection of very-low-abundance molecules in serum by this technology highly unlikely. One should keep in mind that ELISA methodologies, usually used to quantify tumor markers in the circulation, are even more sensitive than Western blotting techniques, allowing direct measurement of analytes at levels as low as 10^{-12} – 10^{-13} mol/liter.

Another methodological artifact that should be kept in mind in SELDI-TOF experiments is the identification of discriminatory peaks (peptides, proteins, or protein fragments) that have originated *ex vivo*. Marshall *et al.* have recently shown that when plasma was left sitting at room temperature for 4 or 8 h, the MALDI-TOF spectra, as recorded by a SELDI-TOF instrument, changed significantly, suggesting that many peptides were generated by proteolytic digestion *ex vivo* (59). The authors attributed this peptide generation to action of specific (serine) proteases, because they could block this effect with serine protease inhibitors. These authors further speculated that the concentration of proteins released into the blood directly from the damaged cells or the changes in important regulatory factors associated with disease are likely to be far too small to be directly detected by MALDI-TOF.

BIOINFORMATIC ARTIFACTS

In virtually every SELDI-TOF experiment published so far, a fraction of the clinical samples are used as a “training set” to derive the interpretation algorithm and the remaining samples used as a “test set.” As correctly pointed out by Qu *et al.* (41), one of the concerns in the construction and use of learning algorithms is the possibility of overfitting the data. It is not

known how robust these algorithms will be when used at different times, or on different sets of clinical samples. One example of demonstrating this possible problem was published by Rogers *et al.* (60). The sensitivity/specificity of a test for discriminating renal cell carcinoma from controls by SELDI-TOF was initially 98–100%. However, when the same procedure was used 10 months later in a new set of patients, the sensitivity dropped to 41%. The authors speculated that this dramatic loss of performance was likely due to sample stability, laser performance, or chip variability. It is thus important to show that algorithms, initially derived with training samples, can still work on different sets of samples and at different times.

It is currently suggested by many authors that m/z ratios <2,000 obtained by SELDI-TOF mass spectrometry should be discarded as noise due to matrix effects (42). However, two of the five discriminatory peaks obtained by Petricoin *et al.* in their ovarian cancer diagnostic protocol with SELDI-TOF had m/z ratios of 534 and 989 (30). In a recent reanalysis of the original raw proteomic data on ovarian cancer, Sorace and Zhan identified peaks that contributed decisively to the discrimination between normal and cancer patients but did not make biological sense (*i.e.* a peak at an m/z ratio of 2.79) (61). These authors raised the possibility for a significant nonbiologic experimental bias between cancer and control groups, casting questions on the validity of the discriminatory peaks with m/z ratios >2,000. Essentially, the same conclusions were reached by Baggerly *et al.* who have also shown that “noise” peaks can achieve perfect classification of normals and cancer patients (62). It is thus mandatory that algorithms used to interpret mass spectrometric data in SELDI-TOF experiments should be carefully reviewed to avoid false conclusions. Indeed, it will be desirable for those working in the field to validate and compare their algorithms and examine if they can come-up with the same discriminatory peaks on the same group of data.

A rather surprising observation relates to two papers on prostate cancer published by the same group (40, 41) (Table II). The same patients were analyzed, and one set of data was generated; this was then examined by two different bioinformatic methods. Surprisingly, the two bioinformatic tools identified different discriminatory peaks. Only two peaks were the same among the nine identified in the first paper and among the twelve identified in the second. Also, one peak that was originally reported in the first study to discriminate between cancer and non-cancer patients was reported to discriminate between healthy controls and patients with benign prostatic hyperplasia in the second study.

In conclusion, it seems that the bioinformatic tools for analyzing SELDI-TOF data need to be carefully validated to avoid artifactual findings and overfitting. Moreover, the reasons for discriminating patients and controls by using peaks within the noise should be further investigated.

TABLE V
Some open questions related to diagnostic SELDI-TOF technology^a

1.	Identity and serum concentration of discriminating molecules not known. Mass spectrometry is a largely qualitative technique. Relationship between peak height and molecule abundance is not linear and could be very complex.
2.	Discriminating peaks identified by different investigators for the same disease are different.
3.	Data not easily reproducible between laboratories, making validation difficult.
4.	Optimal sample preparation for the same disease differs between investigators. Sample handling and preparation may be a critical issue.
5.	Validated serum cancer markers (e.g. PSA, CA125, etc.) that could serve as positive controls are not identified by this technology.
6.	Nonspecific absorption matrices favor extraction of high-abundance proteins/peptides at the expense of low-abundance proteins/peptides. Unknown recovery of "informative" molecules versus "uninformative" molecules. Analytical sensitivity of mass spectrometry in the context of these experiments is not known.
7.	Technique likely measures peptides or other molecules present in high abundance in serum (e.g. mg/liter to g/liter range). Such molecules are unlikely to originate from cancer tissue. More likely, they represent cancer epiphenomena.
8.	No known relationship between discriminatory molecules and cancer biology.

^a This table is reproduced from Ref. 46 with permission from the copyright owners.

INFORMATIVE PEPTIDES IN THE CIRCULATION

Recently, Liotta and colleagues pointed to the possibility that blood contains a vast amount of as yet unutilized and/or unidentified peptides that may have potential as diagnostic biomarkers (63). As indicated earlier, they believe that the tumor-host microenvironment should generate unique signatures in the blood microenvironment. A proposal was then made that the low-molecular-mass region of the blood proteome, which is a mixture of small intact proteins plus fragments of large proteins, represents all classes of proteins and is a treasure trove of diagnostic information largely ignored until now. Because small peptides can be effectively cleared by the kidney, the authors speculate that many of these low-molecular-mass proteins are bound to abundant serum proteins like albumin. This hypothesis needs to be tested experimentally. Experiments to validate this hypothesis should be straightforward. For example, many of these peptides can be characterized by using mass spectrometry. Then, these peptides, tagged with a stable isotope, can be used in recovery experiments to examine if they indeed bind to carrier proteins and what is their lifetime in the circulation. Furthermore, such experiments will reveal their abundance in the circulation and their origin (e.g. which are the parent proteins). Only when these experiments are done will the proposal of using these peptides for diagnostics gain more credence.

NEW INSTRUMENTATION AND EXTERNAL VALIDATION

Liotta and colleagues are now substituting the original CIPHERGEN instrumentation with more sophisticated mass spectrometers of higher resolution. While these instruments do provide improved mass accuracy determination and more complicated spectra, they will not solve the problems outlined in this review regarding diagnostics, because the sample preparation procedures are still based on CIPHERGEN protein chips. Nevertheless, the results of the published and of the newer methods will require careful external validation by different laboratories. Clinical trials are now underway to examine if these methods are robust enough and suitable for clinical use.

Until the external validation data become available, the method should not be used for clinical care.

NORMAL VERSUS ABNORMAL SERUM PROTEOMIC PATTERNS

One explanation for the published data is that the differences in serum proteomic patterns between controls and patients are due to the presence of cancer. Another explanation would be that these differences are not due to the presence of cancer but to a variety of unknown confounding factors. Possible confounders include: sample collection, processing and storage, patient selection and individual habits (e.g. gender, age, ethnicity, exercise, menopausal status, nutritional preferences, drugs, non-cancer diseases, etc.), inappropriate statistical design and/or analysis methods, machine instability, and variable chip performance. The effects of most of these parameters on serum proteomic patterns have not been studied.

Usually, classical tumor markers are evaluated by clinicians by using numerical and easy-to-understand cutoff points (64). All studies using SELDI-TOF technology for diagnostics compare "disease patterns" to "normal patterns." In practice, a normal pattern needs to be generated and used as reference to which the patient pattern will be compared. But how could such a "normal pattern" be generated when the reference group and the testing group are likely to be heterogeneous for the factors described above? It is likely that this "normal pattern" will be influenced by numerous parameters, including diseases different from the one that is being diagnosed. Because SELDI-TOF is a qualitative technique, data interpretation by comparing patterns may prove to be a daunting task.

OUTLOOK

Where should we go from here? In Table V, I summarize some open questions related to this technology. These questions have been posted before (46). Further progress will depend on providing answers after careful experimentation. I would also make some suggestions related to future publications and experiments that need to be done with this technology.

1. Because the technology for positively identifying the discriminatory proteins/peptides is now readily available, future investigations should report the identity of at least some of these peptides and connections made to previous publications associating these proteins with cancer.
2. Future investigations should include internal controls. For example, for common cancers, where classical tumor markers exist (Table I) and their concentrations in the tested samples are known, one goal should be the identification of these biomarkers for validation purposes. Other controls (e.g. exogenously added peptides) should be included and their peak amplitudes used to correct for other peak amplitudes in different experiments.
3. If investigations include “training” and “test” sets, the interpretation algorithm should be tested with an independent series of samples (preferably obtained from another institution) at least 3 months later to validate its robustness.
4. When feasible, it will be useful to apply different bioinformatic algorithms on the same set of data. Do different algorithms produce similar outputs (e.g. the same discriminatory peaks and diagnostic sensitivity/specificity) when the same input data are used?
5. It is imperative that the actual sensitivity of SELDI-TOF mass spectrometry, as it applies to serum analysis, is carefully evaluated. At the same time, it should be investigated if the provided information is quantitative, semiquantitative, or qualitative. Experiments to examine this could include:
 - a. Spiking of serum samples with synthetic peptides 30–50 amino acids long (~molecular mass 3,000–5,000 Da). This could be used to establish the detection limit of these peptides with serum as matrix on various protein chips, and also to establish the quantitative nature of the measurement by constructing calibration curves in a serum matrix.
 - b. Another way of establishing the detection limit of the method would be to select groups of serum samples with known tumor marker concentration (e.g. free PSA of molecular mass around 30,000 Da and α_1 -antichymotrypsin-PSA with molecular mass of 100,000 Da). The quantification of free and total PSA in serum is straightforward using ELISA technologies. Groups of serum samples with free or total PSA concentration, e.g. in the range of 0.5, 2, 10, 50, 200, and 2,000 $\mu\text{g/liter}$, could be run with the SELDI-TOF protocols to establish 1) the detection limit and 2) if the amplitude of the peaks is quantitatively related to the PSA concentration.
6. It will be important to establish in controlled experiments if the discriminatory peptides identified in serum are actually produced *in vivo* or *ex vivo*. Samples from

the same patients should be obtained with or without proteinase inhibitors and processed in various ways, as described by Marshall *et al.* (59), to examine if the discriminatory peaks originate by the action of proteases after the blood is drawn.

7. Studies should be performed to establish the differences in proteomic patterns between plasma and serum and examine the effects of lipemia, length of storage, freeze-thaw cycles, menstrual cycle, age, nutritional status, drug ingestion, sex, race, etc. on such patterns. Without knowledge of the effects of these, and possibly other parameters, on the proteomic patterns generated by MALDI-TOF, the interpretations will be questionable.

CONCLUSIONS

It is true that all classical cancer biomarkers have major shortcomings that preclude their applications for population screening and early diagnosis. The highly promising data generated by SELFI-TOF prompted many to suggest that this technique could be used clinically before the end of this year. However, as indicated above, numerous questions need to be answered before the technology is accepted. There should also be no shortcuts in the validation process of this technology by independent laboratories and agencies. Otherwise, we are running the risk of harming patients who would be misdiagnosed and subjected to unnecessary, invasive, and probably dangerous confirmatory procedures.

As with other medical advances, the ultimate judge of this technology will be time. I sincerely wish that this method will not follow the route of a similar effort originated in the 1980s that suggested cancer diagnosis based on nuclear magnetic resonance profiling of serum samples (65–67).

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. Tel.: 416-586-8443; Fax: 416-586-8628; E-mail: ediamandis@mtsinai.on.ca.

REFERENCES

1. Etzioni, R., Urban, N., Ramsey, S., McIntosh, M., Schwartz, S., Ried, B., Radich, J., Anderson, G., and Hartwell, L. (2003) The case for early detection. *Nature Rev. Cancer* **3**, 243–252
2. Petricoin, E. F., Zoon, K. C., Kohn, E. C., Barrett, J. C., and Liotta, L. A. (2002) Clinical proteomics: Translating bedside promise into bedside reality. *Nature Rev.* **1**, 683–695
3. Srinivas, P. R., Srivastava, S., Hannah, S., and Wright, G. L., Jr. (2001) Proteomics in early detection of cancer. *Clin. Chem.* **47**, 1901–1911
4. Marvin, L. F., Roberts, M. A., and Laurent, B. F. (2003) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in clinical chemistry. *Clin. Chim. Acta* **337**, 11–21
5. Merchant, M., and Weinberger, S. R. (2000) Recent advancements in surface-enhanced laser desorption/ionization time-of-flight-mass spectrometry. *Electrophoresis* **21**, 1164–1167

6. Fung, E. T., and Enderwick, C. (2002) ProteinChip clinical proteomics: Computational challenges and solutions. *BioTechniques* **81**, (Suppl. 34) 40–41
7. Isaac, H. J., Veenstra, T. D., Conrads, T. P., and Felschow, D. (2002) The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem. Biophys. Res. Commun.* **292**, 587–592
8. Isaac, H. J., Conrads, T. P., Prieto, D. A., Tirumalai, R., and Veenstra, T. D. (2003) SELDI-TOF MS for diagnostic proteomics. *Anal. Chem.* **75**, 149A–155A
9. Wulfkühle, J. D., Liotta, L. A., and Petricoin, E. F. (2003) Proteomic applications for the early detection of cancer. *Nature Rev.* **3**, 267–276
10. Petricoin, E. F., Zoon, K. C., Kohn, E. C., Barrett, J. C., and Liotta, L. A. (2002) Clinical proteomics: Translating bedside promise into bedside reality. *Nature Rev. Drug Discovery* **1**, 683–695
11. Pusch, W., Flocco M. T., Leung, S.-M., Thiele, H., and Kostrzewa, M. (2003) Mass spectrometry-based clinical proteomics. *Pharmacogenomics* **4**, 1–14
12. Raj, A. J., Zhang, Z., Rosenzweig, J., Shih, L.-M., Pham, T.-P., Fung, E. T., Sokoll, L. J., and Chan, D. W. (2002) Proteomic approaches to tumor marker discovery. *Arch. Pathol. Lab. Med.* **126**, 1518–1526
13. Chace, D. H., Kalas, T. A., and Naylor, E. W. (2003) Use of tandem mass spectrometry for multianalyte screening of dried blood specimens from newborns. *Clin. Chem.* **49**, 1797–1817
14. Dooley, K. C. (2003) Tandem mass spectrometry in the clinical chemistry laboratory. *Clin. Biochem.* **36**, 471–481
15. Venditti, L. N., Venditti, C. P., Berry, G. T., Kaplan, P. B., Kaye, E. M., Glick, H., and Stanley, C. A. (2003) Newborn screening by tandem mass spectrometry for medium-chain Acyl-CoA dehydrogenase deficiency: A cost-effectiveness analysis. *Pediatrics* **112**, 1005–1015
16. Saugy, M., Cardis, C., Robinson, N., and Schweizer, C. (2000) Test methods: Anabolics. *Baillieres Best Pract. Res. Clin. Endocrinol. Metab.* **14**, 111–133
17. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207
18. Menon, U., and Jacobs, I. (2002) Screening for ovarian cancer. *Best Pract. Res. Clin. Obstet. Gynaecol.* **16**, 469–482
19. Finne, P., Finne, R., and Stenman, U. H. (2001) Neural network analysis of clinicopathological factors in urological disease: A critical evaluation of available techniques. *Brit. J. U. Intl.* **88**, 825–831
20. Stephan, C., Vogel, B., Cammann, H., Lein, M., Klevecka, V., Sinha, P., Kristiansen, G., Schnorr, D., Jung, K., and Leoning, S. A. (2003) An artificial neural network as a tool in risk evaluation of prostate cancer. Indication for biopsy with the PSA range of 2–20 microg/l. *Urologe A.* **42**, 1221–1229
21. Welsh, J. B., Sapinoso, L. M., Kern, S. G., Brown, D. A., Liu, T., Bauskin, A. R., Ward, R. N., Hawkins, N. J., Quinn, D. I., Russell, P. J., Sutherland, R. L., Breit, S. N., Moskaluk, C. A., Frierson, H. F., Jr., and Hampton, G. M. (2003) Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3410–3415
22. Yousef, G. M., Polymeris, M. E., Yacoub, 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
23. Makglara, A., Scorilas, A., Stephan, 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
24. Welsh, J. B., Sapinoso, L. M., Si, A. I., Kern, S. G., Wang-Rodriguez, 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
25. Welsh, J. B., Zarrinkar, P. P., Sapinoso, L. M., Kern, S. G., Behling, C. A., Monk, B. J., Lockhart, D. J., Burger, R. A., and Hampton, G. M. (2001) Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 1176–1181
26. 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
27. Kim, J. H., Skates, S. J., Uede, T., Wong, K. K. K., Schorge, J. O., Feltmate, C. M., Berkowitz, R. S., Cramer, D. W., and Mok, S. C. (2002) Osteopontin as a potential diagnostic biomarker for ovarian cancer. *J. Am. Med. Assoc.* **287**, 1671–1679
28. Diamandis, E. P., and Yousef, G. M. (2002) Human tissue kallikreins: A family of new cancer biomarkers. *Clin. Chem.* **48**, 1198–1205
29. Yousef, G. M., and Diamandis, E. P. (2001) The new human tissue kallikrein gene family: Structure, function and association to disease. *Endocr. Rev.* **22**, 184–204
30. Petricoin III, E. F., 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. (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* **359**, 572–575
31. Powell, K. (2003) Proteomics delivers on promise of cancer biomarkers. *Nat. Med.* **9**, 980
32. Carrette, O., Demalte, I., Scherl, A., Yalkinoglu, O., Corthais, G., Burkhard, P., Hochstrasser, D. F., and Sanchez, J. C. (2003) A panel of cerebrospinal fluid potential biomarkers for the diagnosis of Alzheimer's disease. *Proteomics* **3**, 1486–1494
33. Guillaume, E., Zimmerman, C., Burkhard, P. R., Hochstrasser, D. F., and Sanchez, J.-C. (2003) A potential cerebrospinal fluid and plasmatic marker for the diagnosis of Creutzfeldt-Jacob disease. *Proteomics* **3**, 1495–1499
34. Clarke, W., Silverman, B. C., Zhang, Z., Chan, D. W., Klein, A. S., and Molmenti, E. P. (2003) Characterization of renal allograft rejection by urinary proteomic analysis. *Ann. Surg.* **237**, 660–664
35. Service, R. F. (2003) Recruiting genes, proteins for a revolution in diagnostics. *Science* **300**, 235–239
36. Diamandis, E. P. (2002) Proteomic patterns in serum and identification of ovarian cancer. *Lancet* **360**, 170
37. Diamandis, E. P. (2003) Re: Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* **95**, 489–490. Comment in *J. Natl. Cancer Inst.* 2002 **94**, 1576–1580
38. Burtis, C. A., and Ashwood, E. R. (1994) *Tietz Textbook of Clinical Chemistry*, 2nd Ed., p. 674, W. B. Saunders Co., Philadelphia
39. Petricoin III, E. F., Ornstein, D. K., Paweletz, C. P., Ardekani, A., Hackett, P. S., Hitt, B. A., Valasco, A., Trucco, C., Wiegand, L., Wood, K., Simone, C. V., Leving, P. J., Linehan, W. M., Emmert-Buck, M. R., Steinberg, S. M., Kohn, E. C., and Liotta, L. A. (2002) Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* **94**, 1576–1578
40. Adam, B.-L., Qu, Y., Davies, J. W., Ward, M. D., Clements, M. A., Cazares, L. H., Semmes, O. J., Schellhammer, P. F., Yasui, Y., 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
41. 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 discriminated prostate cancer from non-prostate patients. *Clin. Chem.* **48**, 1835–1843
42. Banez, L. L., Prasanna, P., Sun, L., Ali, A., Zou, Z., Adam, B. L., McLeod, D. J., Moul, J. W., and Srivastava, S. (2003) Diagnostic potential of serum proteomic patterns in prostate cancer. *J. Urol.* **170**, 442–446
43. Lehrer, S., Roboz, J., Ding, H., Zhao, S., Diamond, 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. *Brit. J. U. Intl.* **92**, 223–225
44. Kozak, K. R., Amneus, M. W., Pusey, S. M., Su, F., Luong, M. N., Luong, S. A., Reddy, S. T., and Farias-Eisner, R. (2003) Identification of biomarkers for ovarian cancer using strong anion-exchange Proteinchips: Potential use in diagnosis and prognosis. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12343–12348
45. Service, R. F. (2003) Public projects gear up to chart the protein landscape. *Science* **302**, 1316–1318
46. Diamandis, E. P. (2003) Proteomic patterns in biological fluids: Do they represent the future of cancer diagnostics? *Clin. Chem.* **49**, 1271–1278
47. Wright, G. L., Jr., Cazares, L. H., Leung, S. N., Nasim, S., Adam, B. L., Yip, T. T., Schellhammer, P. F., Gong, L., and Vlahou, A. (1999) Proteinchip® surface enhanced laser desorption/ionization (SELDI) mass spectrometry: a novel protein biochip technology for detection of prostate cancer

- biomarkers in complex protein mixtures. *Prostate Cancer Prostatic Dis.* **2**, 264–276
48. Xiao, Z., Adam, B. L., Cazares, L. H., Clements, M. A., David, J. W., Schellhammer, P. F., Dalmasso, E. A., and Wright, G. L., Jr. (2001) Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. *Cancer Res.* **61**, 6029–6033
 49. Zhang, Z., Bast Jr., R. C., Fung, E. T., Yu, Y., Li, J., Rosenzweig, J., Cameron, B., Rai, A., Sokoll, L. J., Meng, X. Y., Berchuck, A., Van Haften-Day, C., Hacker, N. F., de Bruijn, H. W. A., van der Zee, A. G. J., and Chan, D. W. (2003) A panel of three potential biomarkers discovered from serum proteomic profiling improves sensitivity of CA125 in the detection of early stage ovarian cancer—A multi-institutional study. *Proc. AACR Abstract* 5739
 50. Ye, B., Cramer, D. W., Skates, S. J., Gygi, S. P., Pratomo, V., Fu, L., Hornick, N. K., Licklider, L. J., Schorge, J. O., Berkowitz, R. S., and Mok, S. C. (2003) Haptoglobin-alpha subunit as potential serum biomarker in ovarian cancer: Identification and characterization using proteomic profiling and mass spectrometry. *Clin. Cancer Res.* **9**, 2904–2911
 51. Hlavaty, J. J., Partin, A. W., Shue, M. J., Mangold, L. A., Derby, J., Javier, T., Kelley, S., Steig, A., Briggman, J. W., Haas, G. M., and Wu, Y. J. (2003) Identification and preliminary clinical evaluation of a 50.8-kDa serum marker for prostate cancer. *Urology* **61**, 1261–1265
 52. Cho, W. S. C., Yip, T. T. C., Yip, C., Yip, V., Thulasiraman, V., Ngan, R. K. C., Yip, T.-T., Lau, W.-H., Au, J. S. K., Law, S. C. K., Cheng, W.-W., Ma, V. W. S., and Lim, C. K. P. (2004) Identification of serum amyloid A protein as a potentially useful biomarker to monitor relapse of nasopharyngeal cancer by serum proteomic profiling. *Clin. Cancer Res.* **10**, 43–52
 53. Mueller, W. K., Handschumacher, R., and Wade, M. E. (1971) Serum haptoglobin in patients with ovarian malignancies. *Obstet. Gynecol.* **38**, 427–435
 54. Fish, R. G., Gill, T. S., Adams, M., and Kerby, I. (1984) Serum haptoglobin and alpha-1 glycoprotein as indicators of the effectiveness of cis-diamminedichloroplatinum (CDDP) in ovarian cancer patients—A preliminary report. *Eur. J. Cancer Clin. Oncol.* **20**, 625–630
 55. Fish, R. G., Gill, T. S., Adams, M., Faqeera, F., and Kerby, I. (1984) Changes in serum acute phase proteins in ovarian cancer patients receiving cis-diamminedichloroplatinum (CDDP) infusion therapy. *Clin. Biochem.* **17**, 38–41
 56. Petricoin III, E., and Liotta, L. A. (2003) Counterpoint. The vision for a new diagnostic paradigm. *Clin. Chem.* **49**, 1276–1278
 57. Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* **425**, 737–741
 58. Washburn, M. P., Koller, A., Oshiro, G., Ulaszek, R. R., Plouffe, D., Deciu, C., Winzeler, E., and Yates, J. R., III. (2003) Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3107–3112
 59. Marshall, J., Kupchak, P., Zhu, W., Yantha, J., Vrees, T., Furesz, S., Jacks, K., Smith, C., Kireeva I., Zhang, R., Takahashi M, Stanton, E., and Jackowski G. (2003) Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J. Proteome Res.* **2**, 361–372
 60. Rogers, M. A., Clarke, P., Noble, J., Munro, N. P., Paul, A., Selby, P. J., and Banks, R. E. (2003) Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization (SELDI) and neural-network analysis: Identification of key issues affecting potential clinical utility. *Cancer Res.* **63**, 6971–6983
 61. Sorace, J. M., and Zhan, M. (2003) A data review and re-assessment of ovarian cancer serum proteomic profiling. *B. M. C. Bioinformatics* **4**, 24 (www.biomedcentral.com/1471-2105/4/24)
 62. 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. *Bioinformatics Adv Access*, pp. 1–9, Oxford University Press, Oxford.
 63. Liotta, L. A., Ferrari, M., and Petricoin, E. (2003) Written in blood. *Nature* **425**, 905
 64. Diamandis, E. P. (2002) Tumor markers: Past, present and future, in *Tumor Markers: Physiology, Pathobiology, Technology and Clinical Applications* (Diamandis, E. P., Fritsche, H. A., Lilja, H., Chan, W. C., and Schwartz, M. K., eds) pp. 3–8, AACCC Press, New York
 65. Fossel, E. T., Carr, J. M., and McDonagh, J. (1986) Detection of malignant tumors. Water-suppressed proton nuclear magnetic resonance spectroscopy of plasma. *N. Engl. J. Med.* **315**, 1369–1376
 66. Shulman, R. (1990) NMR—Another cancer-test disappointment. *N. Engl. J. Med.* **322**, 1002–1003. Comment in *N. Engl. J. Med.* 1990 **322**, 949–954; *N. Engl. J. Med.* **322**, 953–958
 67. Engan, T., Krane, J., Klepp, O., and Kvinnsland, S. (1990) Proton nuclear magnetic resonance spectroscopy of plasma from healthy subjects and patients with cancer. *N. Engl. J. Med.* **322**, 949–953. Comment in *N. Engl. J. Med.* **322**, 1002–1003
 68. Diamandis, E. P. (2004) Analysis of serum proteomic patterns for early cancer diagnosis: Drawing attention to potential problems. *J. Natl. Cancer Inst.* **96**, 353–356