Proteomic patterns for early cancer detection

Timothy D. Veenstra, DaRue A. Prieto and Thomas P. Conrads

The advent of proteomics has brought with it the hope of discovering novel biomarkers that can be used to diagnose diseases, predict susceptibility, and monitor progression. Much of this effort has focused on the mass spectral identification of the thousands of proteins that populate complex biosystems such as serum and tissues. A revolutionary approach in proteomic pattern analysis has emerged as an effective method for the early diagnosis of diseases such as ovarian, breast, and prostate cancer. This technology is capable of analyzing hundreds of clinical samples per day and has the potential to be a novel, highly sensitive diagnostic tool for the early detection of diseases, or as a predictor of response to therapy.

Timothy D. Veenstra*
DaRue A. Prieto
Thomas P. Conrads
Laboratory of Proteomics and Analytical Technologies
SAIC-Frederick Inc.
National Cancer Institute at Frederick
Frederick
MD 21702-1201, USA
*e-mail: veenstra@ncifcrf.gov

▼ The rapid development of high-throughput proteomic technologies has brought with it great hope in speeding up the rate at which novel biomarkers for conditions such as early stage cancer are discovered. The positive impact that validated biomarkers can have on the public health cannot be underestimated. Cancer mortality does not arise from a lack of available remedies *per se*, but rather from the diagnosis of such conditions at stages that are too late for remedies to be effective. Fortunately, the ability to characterize proteins within complex biological fluids such as serum, plasma and urine has reached a point where hundreds of species can be identified in a rapid fashion, which brings the greater possibility of identifying needed biomarkers [1,2]. Unfortunately, the sobering reality is that there has not only been a lack of success in the discovery of novel, validated biomarkers, but there have been several spectacular private-sector failures, despite the investment of considerable intellectual and financial resources.

There are different methods by which biofluids such as serum can be analyzed, depending on the goal of the study. Two of these pathways are illustrated in Figure 1. A biomarker-discovery approach aims to identify proteins as they are introduced into the mass spectrometer, regardless of their relative abundance between two cell states. A diagnostic-proteomics approach specifically refers to a high-throughput method that relies on a pattern of peaks to ascertain the condition of the patient from which the biofluid was obtained. A biomarker-discovery approach might use conventional shotgun proteomics, where serum samples are digested into peptides and then individually fractionated and analyzed by liquid chromatography, coupled directly on-line with mass spectrometry (MS). Tandem MS is then used to identify the peptides within each of the mixtures. Subtractive algorithms can be applied to the datasets of peptides identified in the samples to recognize peptides that are either unique to, or more highly abundant in the serum acquired from cancer-affected patients. A popular method to identify aberrantly regulated proteins is fractionation by two-dimensional polyacrylamide gel electrophoresis before tandem MS identification of differentially abundant protein spots [3]. Using a diagnostic proteomics approach, serum samples from healthy and cancer-affected individuals are processed using a protein chip modified with a specific chromatographic surface. After a series of washing steps, matrix is added to the protein spots and the proteomic pattern of each is acquired. Through the use of sophisticated bioinformatic algorithms the source of the sample can be classified as being obtained from a normal patient, cancer-affected patient, or from neither. In contrast to the biomarker-discovery approach, the diagnostic-proteomics approach does not absolutely rely on the actual identification of the protein(s) to diagnose a specific disease state. This approach can, however, continue on to identify these

1359-6446/04/$ – see front matter PII: S1359-6446(04)03246-5

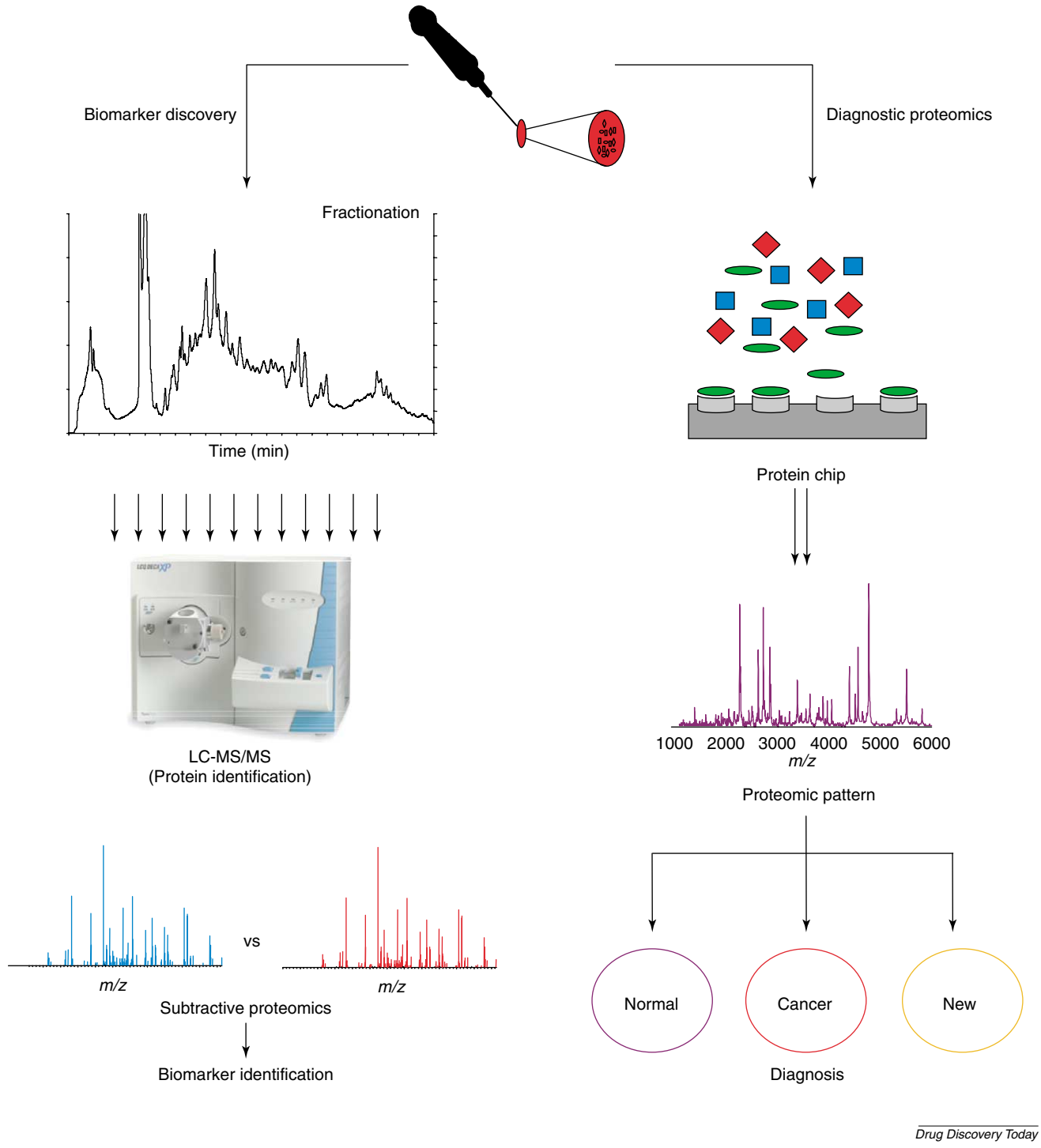

Figure 1. Two commonly used methods for the analysis of biofluids, such as serum, focus on using mass spectrometry (MS) for biomarker discovery or diagnostic proteomics. With biomarker discovery, the focus is on identifying proteins that are unique or highly abundant in samples obtained from patients with specific disease states, compared with healthy, matched controls. This method relies heavily on multidimensional fractionation and tandem MS for protein identification. Using a diagnostic proteomics approach, serum samples from healthy and disease-affected individuals are processed using a protein chip modified with a specific chromatographic surface and a proteomic pattern of the proteins that are retained on the chip is acquired using MS. Bioinformatic algorithms are then applied to the raw data to classify the source of the sample. (i.e. disease-affected, healthy patient or unclassifiable). In contrast to the biomarker discovery approach, the diagnostic proteomic method does not rely absolutely on the actual identification of the protein(s) through which the diagnosis is determined.

differentially abundant peaks, with the further aim of learning more about the biological consequence of the disease state, or to develop an immunoassay designed to measure the abundance of a panel of biomarkers for a specific disease. It is this diagnostic-proteomics approach that has become popularly known as proteomic-pattern analysis.

Proteomic pattern technology

Although a wide variety of different mass analyzers exist for proteomic applications, the overwhelming majority of studies have used surface enhanced laser desorption ionization time-of-flight (SELDI-TOF) MS to acquire proteomic patterns. Three major components constitute the SELDI-TOF–MS: the protein-chip arrays, the mass analyzer, and the data-analysis software [4,5].

Protein-chip arrays

The heart of the SELDI-TOF–MS technology is the protein-chip array [6]. It is this feature that distinguishes it from other MS-based systems used in proteomic research. The protein-chip arrays are available in a variety of different chromatographic surfaces that are designed to retain specific classes of proteins, based on physicochemical properties such as hydrophobicity and charge. The arrays are 10 mm wide and 80 mm long aluminum strips, having eight 2-mm spots comprised of chemically (e.g. anionic, cationic, hydrophobic, hydrophilic, metal ion) or biochemically (e.g. immobilized antibody, receptor, DNA, enzyme) active surfaces. Typically, chemically active surfaces retain whole classes of proteins, whereas biochemically active surfaces are used to couple an antibody or other type of affinity reagent and capture a specific target protein. Whereas the protein chips containing chemically treated surfaces are commercially available, the biochemical surfaces are custom made, by using an open preactivated platform to which a bait molecule can be immobilized. One of the unique features of SELDI-TOF–MS that has made it a popular tool in the analysis of complex biofluids is it ability to analyze very crude samples in an array format, facilitating high-throughput measurements.

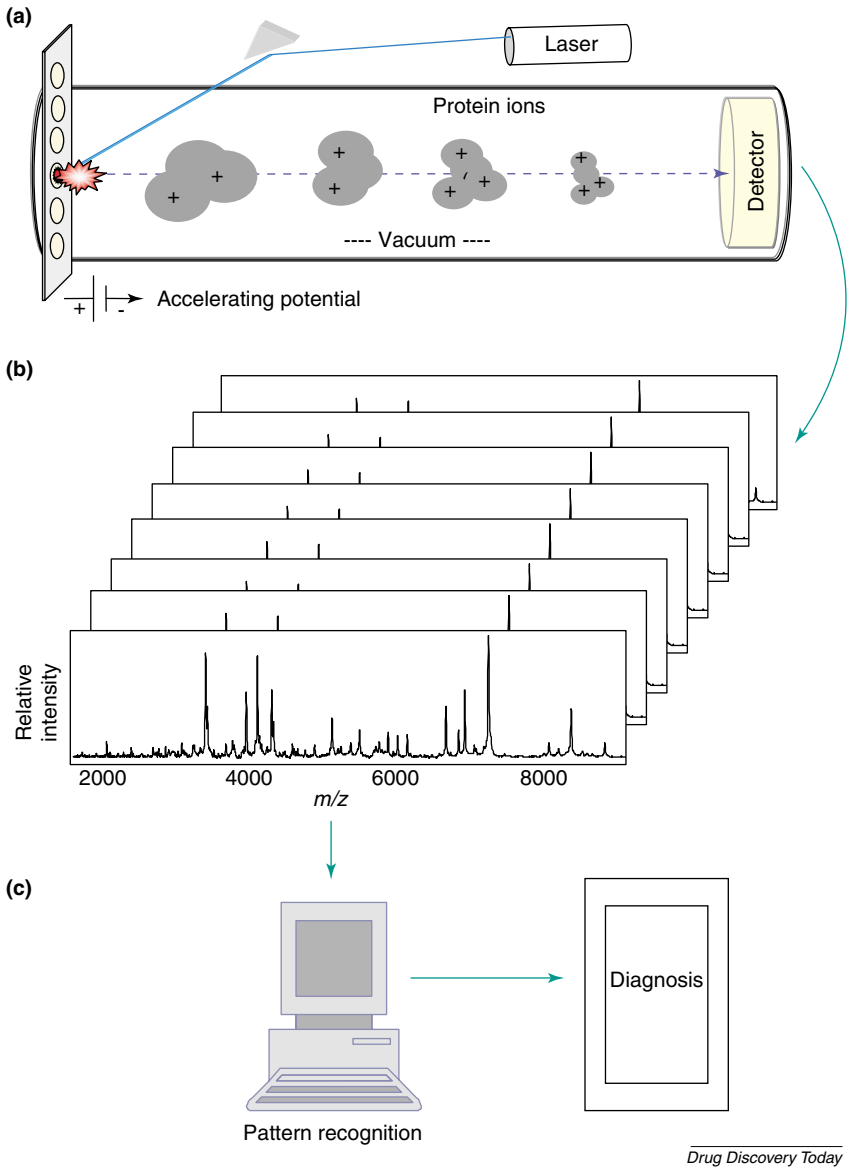

Figure 2. The technology behind the acquisition and analysis of mass spectral patterns for the diagnosis of early-stage ovarian cancer. A serum sample is applied to a protein chip, which is made up of a specific chromatographic surface. **(a)** After a series of washing steps and the application of matrix, a mass spectral image of the species retained on the protein chip surface is acquired. **(b)** For a clinical study, the proteomic patterns of several samples can be acquired in a high-throughput manner. **(c)** Sophisticated bioinformatic software is used to discover peaks within the mass spectral image that enables the pathological condition (i.e. healthy or cancer-affected patient) of the patient from which the serum was acquired to be determined.

The mass analyzer

The most commonly used mass analyzer for acquiring proteomic patterns is a relatively simple TOF–MS equipped

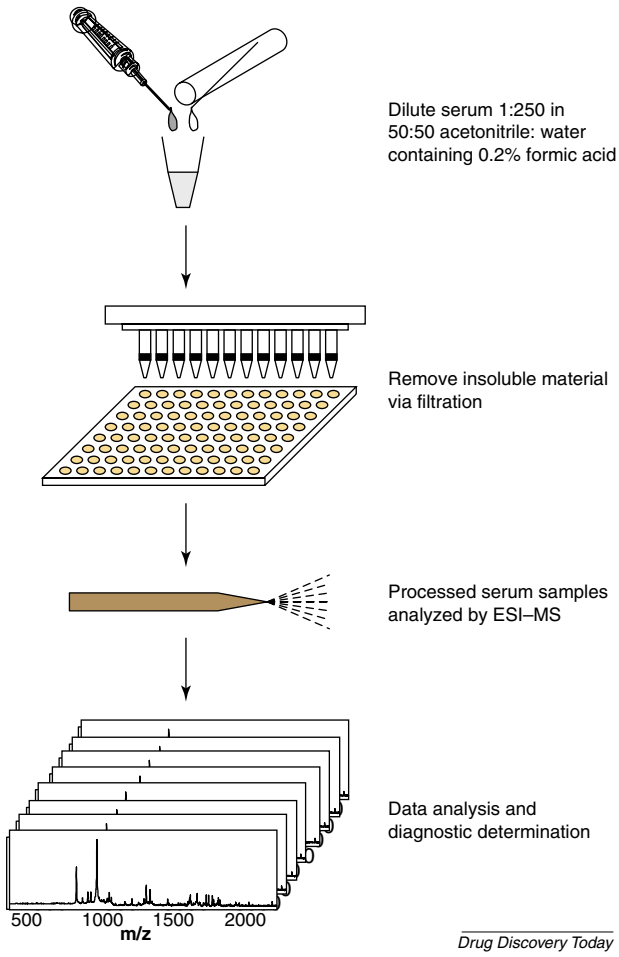

Figure 3. Sample-processing method used to analyze serum samples using the OvaCheck™ diagnostic platform. In this method, raw serum is diluted and filtered to remove particulates. The proteomic pattern of the filtered samples are then acquired using electrospray ionization mass spectrometry. As with SELDI-TOF–MS analysis, bioinformatics is used to discover peaks that enable the source of the serum sample (i.e. a healthy or cancer-affected patient) to be determined.

with a pulsed UV nitrogen laser (Figure 2) [5]. This instrument is known by the commercial name ProteinChip Biomarker System-II (PBS-II) MS. Although SELDI is used to describe the combination of analyzing proteins captured on the ProteinChip surfaces by MS, the gaseous ions are produced by matrix-assisted laser desorption–ionization (MALDI). When a sample is irradiated by a laser, it is transferred into the gas phase (desorption–ionization) and the ionized molecules are accelerated in an electric field, into a field-free region under vacuum (the so-called TOF tube) towards an ion detector. The mass-to-charge ratio (m/z) of the ions is recorded based on the time each species requires to pass through the TOF tube. Compared with high-end spectrometers, the PBS-II TOF–MS has relatively high sensitivity but low resolution and mass accuracy. Based on the primary use of this system, that being the measurement of differences in protein-signal intensities between two or more biological samples in a high-throughput manner, it is appropriate that resolution and mass accuracy be sacrificed for sensitivity.

The resolution, mass accuracy, and lack of tandem MS capabilities of the mass analyzer makes direct protein identification tenuous at best, unless a protein of interest is selectively targeted using an affinity-based surface. So what is the value of the results? The value lies in the ability to obtain spectra from a significant number of samples in a relatively short time period with very little sample preparation or sophisticated chromatography. For example, a single operator can acquire mass spectra of >150 different samples in a single day, or in an automated fashion >500. The analysis of a large number of samples will ideally reveal a pattern of protein signals that are unique to, or overexpressed, in one sample set when compared with a different sample set. The net result is the *m/z* values of protein species that are differentially expressed in different samples.

Proteomic pattern analysis software

The net result of the analysis of a complex proteomic mixture by SELDI-TOF–MS is a low-resolution profile of the protein or peptide species that bound to and were subsequently ionized from the ProteinChip surface. The raw spectra are not particularly visually enticing or dramatic, so why has this technology garnered such fascination and attention? It has been the development and combination of sophisticated bioinformatic algorithms for the analysis of SELDI-TOF–MS data that has led to the potential application of this technology as a major advancement in the diagnosis of cancer and other diseases. There are several different types of bioinformatic algorithms, such as single classification trees, neural nets, genetic algorithms, and random forest algorithms, have been applied to enable SELDI-TOF–MS data to be investigated as a diagnostic technology [7–10]. Although they function in different manners, these algorithms share a common goal: to construct a classifier and discover peak intensities most likely to be responsible for segregating classes of samples. Since its inception, SELDI-TOF–MS has been used to develop diagnostic platforms for several different cancers, including breast [11], prostate [12], head and neck [13], gastrointestinal [14], pancreatic [15] and ovarian [8].

Technological improvements

High-resolution proteomic patterns

The original proteomic pattern studies were conducted using the PBS-II MS platform, and this instrument is still

the most widely used technology for doing these types of studies. This MS instrument is well suited for many biological science laboratories because it is easy to use and does not require substantial training in the field of MS or analytical chemistry to obtain worthwhile data. In a comparison with other types of TOF–MS instruments, the PBS-II provides reasonably high sensitivity; however, it suffers from poor resolution and mass accuracy. Regardless of these deficiencies, the diagnostic results reported from studies conducted using the PBS-II have been impressive.

As with any developing area of research, there is always room for improvement in the overall technology. One of the first major advances was the use of a high-resolution hybrid quadrupole TOF (QqTOF) MS [16] fitted with a SELDI ion source to acquire proteomic patterns from serum. A recent study was designed to determine whether there is any diagnostic advantage provided by acquiring the proteomic patterns of serum samples using a high-resolution, high mass accuracy MS instrument [17]. In this study, 248 serum samples from healthy and ovarian-cancer afflicted patients were analyzed using both a high-resolution hybrid quadrupole TOF (QqTOF) MS fitted with a SELDI ion source and a PBS-II instrument. An identical set of samples was analyzed on the exact same ProteinChip surface, thereby eliminating all experimental variability outside the use of two different instruments. Different combinations of bioinformatic heuristic parameters were used to generate 108 diagnostic models using the data acquired from the two distinct mass spectrometers. These parameters included the similarity space of likeness for cluster classification, the feature-set size of random *m/z* values whose combined intensities comprise each pattern, and the learning rate in training of the genetic algorithm. The derived diagnostic models were validated using blinded serum sample spectra obtained from 37 unaffected women and 40 women with ovarian cancer. The diagnostic models generated from mass spectra acquired using the higher-resolution Qq-TOF MS were statistically superior, not only in testing but also in validation to those acquired on the PBS-II.

Several diagnostic models were found that were able to correctly classify the source (i.e. from healthy or ovarian cancer-affected patients) with 100% sensitivity and specificity. Each of these models was generated with data acquired on the Qq-TOF–MS; no models with both 100% sensitivity and specificity could be found using the PBS-II data [17]. Quite importantly, not only were the 81 cases of stage II, III and IV ovarian cancer correctly classified, but so were all 22 cases of women with stage I ovarian cancer.

Magnetic particle sample preparation

In one of the most recent developments in proteomic pattern technology, serum peptides are captured and concentrated using reversed-phase (RP) batch processing in a magnetic particle-based format and analyzed by MALDI- TOF–MS [18]. In this sample handling procedure, a suspension of magnetic beads that are coated with reversed phase chemistry is mixed with 50 µl of serum and the beads are pulled to the side of the tube by magnetic force. After removal of the supernatant, the beads are subjected to a series of washing steps, after which an elution solvent containing 50% acetonitrile is added to remove the bound serum peptides from the surface-modified magnetic beads. At this stage, the supernatant (which contains the serum peptides of interest) is carefully transferred to a second tube and matrix solution is added to the eluate and mixed. An aliquot of this mixture is transferred to the MALDI target and the proteomic patterns of the various samples acquired. This sample preparation procedure has significant advantages over the SELDI-based methods in that it enables serum proteome patterns to be acquired on a myriad of different TOF-MS instruments. Presently, there are only two different types of spectrometers that accept a SELDI interface; the PBS-II (manufactured by Ciphergen Biosystems; http://www.ciphergen.com) and the QqTOF (manufactured by Applied Biosystems; http://www. appliedbiosystems.com). The sample-preparation technique presented here enables proteomic patterns to be acquired on any MS capable of MALDI. These types of ion sources are ubiquitous in almost every MS laboratory or core facility and when coupled with TOF instruments are capable of producing spectra with resolution in excess of 10,000, mass accuracy within 50 ppm, and sensitivity in the femtomolar range. The sample preparation step is automatable and can easily prepare hundreds of samples per day.

This group applied this new sample preparation technology to a pilot study with the goal to be able to distinguish patients with glioblastoma (GBM) from controls (i.e. no evidence of cancer). There are few clinically useful serum markers for GBM, however, it is the most common and most malignant brain tumor to arise in adults [19]. Serum samples from 34 GBM patients and from 22 healthy volunteers were processed as described above and analyzed using an Ultraflex MALDI-TOF–MS manufactured by Bruker (http://www. bdal.com). Each spectrum contained >400 distinct peptide peaks, which were unambiguously detected in each sample. After being aligned through their m/z values ('binning'), almost 1700 unique peaks were found in the 56 cumulative spectra. A statistical difference (p <0.05) between GBM and control cases was found in 274 peaks. These peaks were then used to cross validate the ability to discriminate the classes. Two classes were created by using 55 out of the 56 samples as a training set and using the 56th sample as a test set to verify whether it would be

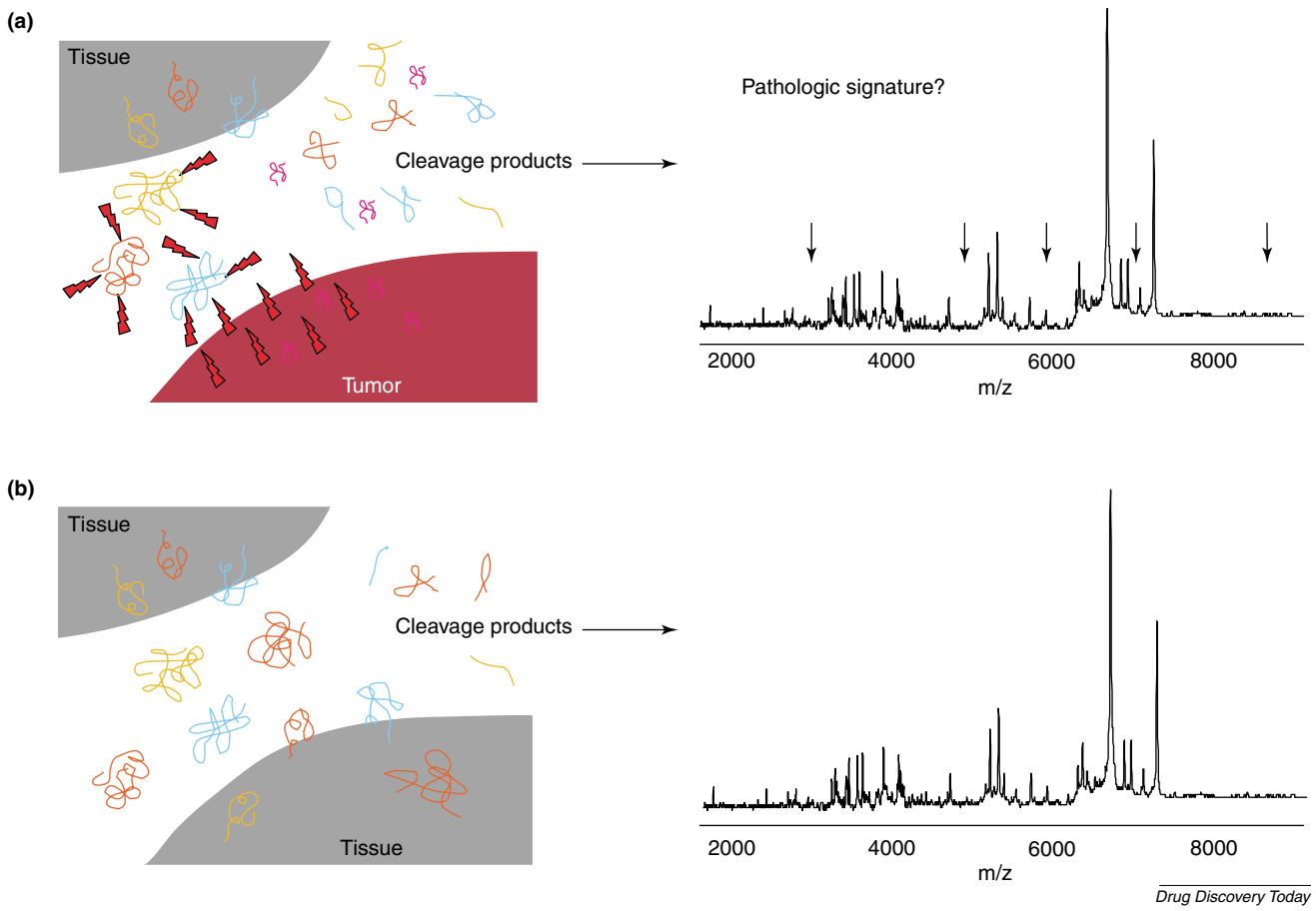

Figure 4. Hypothesis of how a tumor generates signature proteins within the serum of cancer-affected patients. **(a)** Tumors secret aberrant protein products or produce distinct cleavage products of proteins already within the circulatory system through the action of tumor-specific metalloproteinases (shown as red lightening bolts). It is the presence of these cleavage products and tumor-secreted proteins that give rise to a pathological signature within the mass spectra of serum acquired from cancer-affected individuals. **(b)** In the absence of a tumor, these cleavage products or tumor-secreted proteins are not present within the blood and therefore the pathological signature indicative of cancer is absent in the mass spectrum.

correctly classified. The process was repeated 56 times, that is, until all samples had been used as a test set. This analysis was able to correctly classify 53 out of 55 (96.4%) samples. Two controls incorrectly classified (3.6%) as GBM samples, whereas the remaining sample was not classified. Although this study was conducted using a limited number of samples, nonetheless, it is the first of its kind to employ a straight MALDI-based approach for the acquisition of serum proteomic patterns for the purpose of disease diagnostics.

OvaCheck™

Recently, a new approach for acquiring serum proteomic patterns for diagnosing ovarian cancer has attracted a great deal of attention in the public media. This technique, which is being licensed under the trademark OvaCheck™ [20], uses electrospray ionization (ESI) instead of MALDI or SELDI, and does not require the use of protein chips to prepare the serum sample for mass spectral analysis. The sample preparation for the samples is quite minimal, requiring the serum to be diluted 1:250 in a 50:50 mixture of acetonitrile and water, containing 0.2% formic acid and then removing the insoluble particules via filtration (Figure 3). The serum samples are collected into a 96-well plate that is placed into the interface of an automated ESI interface, the Nanomate 100 MS (manufactured by Advion Biosciences; http://www.advion.com) [21]. As opposed to the SELDI-based approach that is based on MALDI, the Nanomate 100 uses ESI to create positively charged gas-phase ions that can be measured by MS.

Although the published reports of a serum-based test for various cancers using the SELDI-TOF–MS platform have been around for over two years (and in development for over four years), there are no peer-reviewed published reports using the OvaCheck™ platform. There have been

some public disclosures of this technology at scientific meetings, including a recent poster presented at the Society of Gynecological Investigation conference in Houston, TX on March 25, 2004. The poster presented results that were 97% sensitive and 94% specific in validation for the diagnosis of ovarian cancer. Why then, without a significant track record or available peer-reviewed reports to analyze, has this test received such notoriety? There are two reasons. First, it is very similar to the SELDI-TOF–MS based diagnostic test. Obviously, since the first report demonstrating the enormous potential of SELDI-TOF–MS for the diagnosis of early stage ovarian cancer, there has been a huge influx of investigators who are evaluating the technology on serum, plasma, and other biofluid specimens to determine the validity of this technology on ovarian and other cancers. Second, the marketing and licensing strategy by Correlogic Systems (http://www.correlogic.com) who hold the patent and licensing rights to OvaCheck™, has generated substantial publicity for this test, especially in the public arena. Not only has OvaCheck™ been profiled on NBC's Today Show, there are recent reports describing it in several major daily newspapers including the Philadephia Inquirer (http://www.correlogic.com/inquirer.pdf) and the Miami Herald (http://www.miami.com/mld/miamiherald/business/national/8430195.html).

How has the scientific community reacted to the publicity and claims generated around OvaCheck™? In two words: politely unfavorable. The Society of Gynecologic Oncologists recently issued the following statement; 'In the opinion of the SGO, more research is needed to validate the test's effectiveness before offering it to the public' (http:// www.sgo.org/images/pdfs/policy/OvaCheck_statement.pdf).The Food and Drug Administration (FDA) had a negative response to Correlogic System's intent to license OvaCheck™ to major commercial laboratories as a test for diagnosing ovarian cancer. In a publicly available letter (http://www. fda.gov/cdrh/oivd/letters/021804-correlogic.html) sent to Correlogic Systems, the FDA stated that they had no record of OvaCheck™ being subjected to a premarket review. Although the FDA and SGO are very supportive of pursuing the proper validation of OvaCheck™, and recognize the need for better diagnostic tools for ovarian cancer, the virtues of this test have been prematurely celebrated in the public forum. OvaCheck™, as well as all of the proteomic pattern research studies conducted using SELDI or MALDI platforms, need to pass rigorous validation studies before any of these technologies are ready for clinical adoption.

Wherein lies the value of proteomic patterns?

Although having exploded onto the clinical and research community with great promise and excitement, the value and future of proteomic pattern technology in diagnostic medicine is hotly debated, often fueled by what the ultimate purpose and endpoint of analyzing biofluids using this technology is. All of the early studies in this field used the collection of peaks observed in the spectra of samples from control and diseased patients to find a fingerprint that is indicative of the pathophysiological state of the patient. Few studies proceeded to identify the discriminatory peak and the overriding theme was that 'the diagnostic power is within the pattern'. That is, if the collection of discriminatory peaks could correctly diagnose the pathophysiological state of the patient with a high degree of sensitivity and specificity, what is the value to the patient (or their doctor) in knowing the identity of the discriminating peaks? This premise underlies one facet of the power of this technology within a clinical setting. As hundreds of clinical samples can be run in a single day on a single SELDI-TOF–MS instrument, it is not outside of reality to envision centers where thousands of these samples could be analyzed per day.

Another view is that it is important to identify the peaks discriminating healthy and disease samples as these could be classical biomarkers to which an immunoassay can be designed to measure in a clinical test. In addition, the identification of these peaks might lead to further understanding of tumor progression or recurrence in the patient. In this research area, peaks that are found to be differentially abundant based on the SELDI-TOF–MS profiles comparing samples from control and disease cases, are targeted for identification. This school of thought likens this technology to biomarker discovery, in which the aim is to conclusively identify a protein that is specific to a disease state. Unfortunately, the normal SELDI-TOF–MS procedure is limited in the view it provides of complex mixtures, such as serum. This limited view is reflected by the identity of many of the protein signals that have been shown to be differentially abundant and therefore an indicator of disease state. For example, haptoglobin was identified as a marker for renal cancer and ovarian cancer, and serum amyloid A was identified as markers for renal cancer [22,23]. Elevated haptogloblin levels were originally identified >30 years ago by classical techniques as putative tumor marker for ovarian cancer using classical techniques [24]. This protein was never used for clinical diagnosis because of its low sensitivity and specificity. It has been known for almost 30 years that increased levels of serum amyloid A result from some forms of cancer as well as acute bacterial and viral infections [25]. Presently, there does not remain an instance in which a single biomarker with high sensitivity and specificity for a particular disease has been discovered using SELDI-TOF–MS. Although some groups have

used chromatographic separation of complex samples before spotting on ProteinChips [26], this limited separation, combined with the low resolution and dynamic range afforded by the PBS-II spectrometer, affords a restricted capability for observing a complex proteome. Studies attempting to identify a classical biomarker require more-sophisticated separations and higher-end MS technologies (i.e. FT-MS and QqTOF) that have a greater capacity for observing species within extremely complex mixtures.

What is the source of the diagnostic information in proteomic patterns?

Although proteomic pattern technology has shown considerable promise, it is not without its critics. The criticism revolves around the undeniable fact that nobody knows for certain how this combination of MS and bioinformatics is able to determine the pathological condition of a patient. There is no conclusive evidence for the source of the diagnostic information (i.e. classifying peaks) contained within the serum profiles. One hypothesis is that, because serum is constantly perfusing tissues throughout the body, its composition is influenced by the various types of tissues it comes into contact with (Figure 4) [27]. These influences manifest not only in the secretion and shedding of proteins from cells into the bloodstream but also in the processing of common highly abundant circulatory proteins by proteins originating from various cells and tissues. Because tumors are pathologically distinct, their influence on serum would differ from normal, healthy tissues. This distinct influence could be in the form of cancer-specific proteins or the action of metalloproteinases on highly abundant circulatory proteins. It is these tumor-specific proteins that are hypothesized to constitute the distinctive peaks allowing serum from cancer-affected individuals to be segregated from that acquired from healthy individuals.

The second hypothesis is that the peaks that allow this discrimination do not originate from the tumor [28]. These distinguishing peaks originate from proteins that are epiphenomena of cancer and are produced by other organs responding to the presence of cancer. They might also be related to a generalized condition of the cancer patient, such as infection, malnutrition, or general ill health. If this is the case, then it will be very difficult to distinguish different cancers from one another or possibly even different disease states from cancer. Most of the studies using SELDI technology have been aimed at distinguishing normal patients from those suffering from a particular disease state. One of the future steps will be to determine how effective this technology is at specifically distinguishing disease states from one another. It remains to be seen whether these molecules could indeed collectively constitute specific biomarkers for cancer, in view of the fact that cancer epiphenomena are not disease specific. Unfortunately, at present there is not enough hard evidence to determine which of these two hypotheses more accurately describes reality.

Reproducibility of SELDI-TOF–MS

An effective diagnostic test needs to be reproducible and provide identical results every time it is used, regardless of the laboratory it is being conducted in. Unfortunately, the lack of reproducibility seen in SELDI-TOF–MS studies has been a major criticism and will be one of the major hurdles that this technology needs to clear before it can be a useful clinical tool. A case in point is the comparison of two studies using SELDI-TOF–MS to develop diagnostic markers for prostate cancer [12,29]. Each of these studies used the same type of protein chip surface and obtained their MS data using the same type of spectrometer. Both of these studies could diagnose prostate cancer with comparable sensitivity and specificity, however, there was a lack of similarity in the peaks selected as diagnostic markers for discriminating serum obtained from prostate-affected and control individuals. Only two peaks, out of nine and twelve discovered in the respective studies, were common to both studies (i.e. *m/z* 7024 and 9656), even though these studies used the same sample analytical platforms.

Although the true causes of the irreproducibility between these two studies is uncertain, there are two obvious possibilities. First, the spectral patterns acquired by SELDI-TOF–MS are very sensitive to every step in the entire analytical process originating from the time the serum is collected until the spectrum is acquired. Unfortunately, no universal standard protocol exists for the collection, storage, transport and aliquoting of serum samples exists. In addition, there is no universal quality control and assurance method for the processing of samples or instrumental parameters for the acquisition of the spectra. Second, the application of different bioinformatic algorithms to datasets, as was the case in these two studies, will invariably result in the identification of different sets of features within the mass spectra that are used to discriminate cancer cases from controls. The chances that two different groups would find the same discriminating peaks using different instruments and computer algorithms would be extremely low. Unfortunately, neither of these explanations alleviates the problem of laboratory-to-laboratory irreproducibility when using SELDI-TOF–MS. The ability to obtain and analyze these data in a reproducible manner will be one of the determining factors that dictate whether this diagnostic method makes an impact in the clinical laboratory someday.

Conclusions

Proteomic pattern technology is still in its infancy and, although its potential as a diagnostic tool is great, much still needs to be learned. The entire analytical process requires optimization, beginning with sample collection and ending with data analysis. The National Cancer Institute (http:// www.cancer.gov) has charged a clinical reference laboratory to conduct such an evaluation. This laboratory's mission is to evaluate every stage of this technology with the goal to garner FDA approval for the use of proteomic patterns to diagnose ovarian cancer (or its recurrence) in high-risk populations. Until further validation studies are performed this technology should not be considered ready for clinical application but is still a viable research tool.

Although many critics still abound, the fact that the diagnostic models generated from proteomic patterns continue to provide highly sensitive and specific results in testing and blind validation studies, cannot be ignored. There continues to be an increase in the number of laboratories worldwide that have adapted this technology as part of their discovery proteomics efforts. Ultimately, it will be a combination of these individual efforts that could determine whether proteomic pattern analysis will have the kind of impact on diagnostic medicine that one can hope for.

Acknowledgements

This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. NO1-CO-12400.

By acceptance of this article, the publisher or recipient acknowledges the right of the U.S. Government to retain a nonexclusive, royalty-free license and to any copyright covering the article. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1 Chan, K.C. *et al.* (2004) Analysis of the human serum proteome. *Clin. Prot* 1, 101–226
2 Anderson, N.L. *et al.* (2004) An expanded non-redundant human plasma proteome list developed by combination of four separate sources. *Mol. Cell. Proteomics* 3, 311–326
3 Anderson, N.L. and Anderson, N.G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 1, 845–867
4 Hutchens, T.W. and Yip, T.T. (1993) New desorption strategies for the mass spectrometric analysis of macromolecules. *Rapid Commun. Mass Spectrom.* 7, 576–580
`5 Issaq, H.J. *et al.* (2003) The surface enhanced laser desorption ionization time of flight mass spectrometric approach to diagnostic proteomics: protein profiling and biomarker detection. *Anal. Chem.* 75, 148–155
6 Weinberger, S.R. *et al.* (2002) Current achievements using ProteinChip array technology. *Curr. Opin. Chem. Biol.* 6, 86–91
7 Izmirlian, G. (2004) Application of the random forest classification algorithm to a SELDI-TOF proteomics study in the setting of a cancer prevention trial. *Ann. N. Y. Acad. Sci.* 1020, 154–174
8 Petricoin, E.F. *et al.* (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 16, 572–577
9 Holland, J.H. ed. (1994) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, (3rd edn), MIT Press Cambridge, Massachusetts, U. S. A.
10 Kohonen, T. (1990) The self-organizing map. *Proc. IEEE* 78, 1464–1480
11 Laronga, C. *et al.* (2003) SELDI-TOF serum profiling for prognostic and diagnostic classification of breast cancers. *Dis. Markers* 19, 229–238
12 Qu, Y. *et al.* (2002) Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. *Clin. Chem.* 48, 1835–1843
13 Wadsworth, J.T. *et al.* (2004) Identification of patients with head and neck cancer using serum protein profiles. *Arch. Otolaryngol. Head Neck Surg.* 130, 98–104
14 Fels, L.M. *et al.* (2003) Proteome analysis for the identification of tumor-associated biomarkers in gastrointestinal cancer. *Dig. Dis.* 21, 292–298
15 Rosty, C. *et al.* (2002) Identification of hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein I as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology. *Cancer Res.* 62, 1868–1875
16 Chernushevich, I.V. *et al.* (2001) An introduction to quadrupole-time-of-flight mass spectrometry. *J. Mass Spectrom.* 36, 849–865
17 Conrads, T.P. *et al.* (2004) High-resolution serum proteomic features for ovarian cancer detection. *Endocr. Relat. Cancer* 11, 163–178
18 Villanueva, J. *et al.* (2004) Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal. Chem.* 15, 1560–1570
19 Holland, E.C. (2001) Gliomagenesis: genetic alterations and mouse models. *Nat. Rev. Genetics* 2, 120–129
20 Fishman, D.A. *et al.* (2004) *High-throughput multidimensional mass spectrometry analysis for the detection of early stage epithelial ovarian cancer: a serum test for ovarian cancer*. Society of Gynecological Investigation Houston, TX, U. S. A.
21 Van Pelt, C.K. *et al.* (2003) A fully automated nanoelectrospray tandem mass spectrometric method for analysis of Caco-2 samples. *Rapid Commun. Mass Spectrom.* 17, 1573–1578
22 Tolson, J. *et al.* (2004) Serum protein profiling by SELDI mass spectrometry: detection of multiple variants of serum amyloid alpha in renal cancer patients. *Lab. Invest.* 84, 845–856
23 Ye, B. *et al.* (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
24 Mueller, W.K. *et al.* (1971) Serum haptoglobin in patients with ovarian malignancies. *Obstet. Gynecol.* 38, 427–435
25 Benson, M.D. *et al.* (1977) Kinetics of serum amyloid protein A in casein-induced murine amyloidosis. *J. Clin. Invest.* 59, 412–417
26 Koopmann, J. *et al.* (2004) Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. *Clin. Cancer Res.* 10, 860–868
27 Petricoin, E.F. and Liotta, L.A. (2004) SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr. Opin. Biotechnol.* 15, 24–30
28 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
29 Adam, B.L. *et al.* (2002) Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.* 62, 3609–3614