



LABORATORY MEDICINE IN THE SCOPE OF PROTEOMICS AND GENOMICS

Urszula Demkow

Dept. of Laboratory Diagnostics and Clinical Immunology of Developmental Age, Medical University of Warsaw

Correspondence

urszula.demkow@litewska.edu.pl

Abstract

Advances in technology, especially in molecular biology, allow for a fast expansion of diagnostic methods in routine clinical practice. New proteomics and genomics technologies could be used for disease specific biomarker discovery and to monitor patient response to the therapy. Genomics and proteomics may also help to establish new, molecular classification of the disease. Applying genomic and proteomic methods to body fluids (serum, cerebrospinal fluid, urine, etc) and tissue extracts would place valuable objective analytical power in the hands of the clinician however validation of those methods is an important issue. The rapid expansion of the diagnostic tools based on developments in proteomic and genomic technologies can be fundamental for the development of personalized medicine.

INTRODUCTION

Nowadays in modern clinical laboratories chemical, hematologic and immuno-chemical parameters can be efficiently measured using automated equipment. Advances in technology with improved understanding of molecular pathology, allow for a fast expansion of diagnostic methods and multiplication of the parameters measured in the laboratory. The development of miniature laboratory chips will reduce the clinical sample and reagent volumes and allow more data to be obtained in a shorter period of time (1). Currently laboratory tests are used for diagnosis and monitoring of disease activity or therapy. The future tests will focus on predisposition testing, targeted monitoring, and prevention of diseases through nutrition, lifestyle and drug therapy (2). Modern diagnostics is changing from isolated medicine to personalized medicine (2). This approach may create a greater opportunity to prevent diseases. The rapid expansion of the diagnostic tools can be attributed to developments in proteomic and genomic technologies.

GENOMICS

The human genome sequence was described in 2001, creating an opportunity for personalized medicine (3). Genomics is a scientific discipline that characterizes the complete genome of an organism. Its primary approach is to determine the entire sequence of DNA and the relations between different parts of the genome. Transcriptomics is the part of the genomics that describes mRNAs encoding individual proteins.

Genomics-based devices have the potential to become first line tools to identify patients at risk for developing certain diseases or predict unusual reactions to certain drugs. This knowledge could be widely used in clinical medicine. Previously, genes for monogenic or Mendelian disorders were easily discovered (4, 5). Genomic's tools may allow to find genetic background for complex, multi-factorial disorders like autoimmune diseases, cancer, cardiovascular disorders or even schizophrenia (2, 4). The location of numerous DNA markers can be compared in samples from patients and healthy controls, and calculated statistical difference can confirm or deny the linkage (4). Such investigations has been undertaken in cancer or schizophrenia patients (2, 4).

An array of high-throughput technologies dedicated to genomics research are used to characterize the biological function of genes and genomes (6, 7, 8). For example PCR and DNA sequencing are frequently used to search for genetic variants associated with the development of a disease (8). Innovative, high-throughput microfluidic systems

and automating of the processes involved in genomics research including: extraction and purification of nucleic acids, set up of PCR reaction, automation of RT-PCR as well as the separation and direct detection of DNA and RNA, are on the development too (9, 10, 11). These advanced technologies combine microelectronics with molecular biology tools. The study of gene expression is enabled through microarrays and qRT-PCR techniques that measure RNA levels related to specific gene transcription processes, as well as other RNA-mediated processes (12).

A gene microarray (gene chip) is a miniaturized slide that carries numerous probes of nucleic acids, which are arranged in a grid pattern on the chip. Microarrays are useful because of their small size and because they can examine a very large number of genes. Microarrays can rapidly provide a detailed view of the simultaneous expression of all the genes (around 30 000) in an entire genome, and provide new insights into gene function, disease pathology and classification, and drug development (13). The main challenge in microarray technology results from the technical complexity of the process and the large amounts of data generated in the experiment (12, 13).

Advanced array technologies include genome sequencing, genotyping, transcriptome analysis but also protein analysis, functional cell microarrays and tissue microarrays (14). Genomic microarrays enable to measure the expression of genes under different medical conditions. A small amount of body fluid has the ability to look at 30 000 genes (12, 13, 14). Genomic microarrays may become useful in early screening of diseases such as lung cancer, which usually are not diagnosed until they are advanced and less treatable (15). In the future, microarray-generated data may help clinicians with earlier cancer classification and diagnosis. The emergence of high-throughput microarray profiling allowed the rapid and economical assaying of thousands of gene's expression levels. Whenever genetic material from a patient's sample is placed on the DNA chip, the probes react (6, 8, 13, 14). Those reactions can be detected and used to screen for the presence of particular genetic sequences, such as those related to diseases or can predict unusual reaction to certain medications (6, 8). For example cytochrome P450 genotyping tests detects variations in a gene that affects liver metabolism of certain drugs, such as antidepressants, antipsychotics and selected chemotherapeutic agents (16). Then the drug's dosage for an individual patient can be adjusted.

Microarray technologies under development include search for genotypes related to different types of cancer, allergy, autoimmune diseases, cystic fibrosis and genes having impact on drug metabolism and response (6, 16). Specific medical applications will also include screening for inherited genetic disorders (such as amyotrophic lateral sclerosis and muscular dystrophy), combating diseases of the nervous system (such as Alzheimer's disease, Parkinson's disease and new variant Creutzfeldt-Jakob disease), cardiovascular diseases, rare diseases, as well as for detection of infectious agents (14, 15, 16, 17). For example, cystic fibrosis genomic tests find genetic variations in one of the genes that causes cystic fibrosis - the most common fatal genetic disease. Those tests help to diagnose cystic fibrosis in children and identify adults carrying the defective gene (18). Other projects being studied are the diagnostic gene expression microarrays for allergy. Microarrays are currently used to search for the antigen HLA B27 in rheumatic disease patients and HLA DQ2, DQ8 in celiac disease (20).

New methods for array-based resequencing include oligonucleotide fingerprinting, iFRET technology to detect hybridization, nanoscale hybridizations in nanowells, and MALDI mass spectrometry to detect oligonucleotide composition (21). A high throughput resequencing device may become a very competitive tool, which would create possibilities for genome-wide sequence evaluation of patient biological material as a routine procedure (21, 22). Moreover monitoring the activity of a genome by measuring mRNA expression levels provides important biological insights.

Another recently introduced method is RNA interference (RNAi), an effective mechanism for selective inhibition of gene expression which, has become the preferred method for inhibiting expression of targeted genes (23). As well as functional genomics applications, it also shows tremendous potential for diagnostics and therapeutics. RNAi libraries covering the entire genome are being developed to secure a functional validation of gene targets.

Recent clinical studies demonstrate that expression analysis of large gene sets can identify molecular profiles correlated to disease states, which may be used for the construction of diagnostic tools (24, 25). However, frequently, the molecular characterization of clinical sample is limited by the available volume of sample. Amplification of the starting material or of the signal to be detected, or frequently miniaturization of the method can help to overcome this problem (24, 25). Novel means are also required to measure gene expression with allele-specific and splice-variant-specific profiles (24, 25). The above mentioned technologies will likely be a big step from bench to bedside. However, the performance of a microarray-based method has to be properly evaluated before transfer to the clinic. These methods need to be highly automated, miniaturized, easy to use and inexpensive as well as must assure the ability to analyze DNA sequences accurately and rapidly. The expansion of microarray-based technologies will probably have a major impact on the evaluation of laboratory tests as diagnostic tools. However, advanced techniques

such as microarrays leave many opportunities for errors. The lack of standardization for naming and identifying the genes used on different DNA microarray platforms could cause potential errors. Before microarrays can be consistently and reliably used in clinical practice, and in decision making, standards, quality control, and interpretation issues need to be settled (26).

PROTEOMICS

Although genomic data may discover novel information on the pathogenesis of numerous disorders, finding useful diagnostic laboratory markers may be within the scope of proteomics (27). Gene expression does not consistently correlate with protein expressions, and cannot identify post-transcriptional and post-translational modifications, major modulators of protein function (4, 28). In the "post-genomic" era, the progress is towards examining proteins as the main effectors of physiological functions (4). Detailed characterization of the proteome (total body proteins) is a major goal of proteomics, that analyses disease mechanisms by examining changes in the patterns of proteins in patient's body fluids and tissues. The analysis of complex protein mixtures such as serum, other body fluids or tissues by profiling hundreds of proteins in the same time, creates pattern of response characteristic for various cellular states or disease conditions. Detailed proteome analysis has become more realistic today with the high-resolution mass spectrometers capable of faster sequencing in a high-throughput fashion and with the emergence of new techniques such as microarrays. A promising area is the application of advanced mass spectrometric and other quantitative proteomic methodologies to laboratory diagnostics (28, 29, 30). The major proteomic projects of the last decade have shaped proteome-wide sequencing, mapping, and analysis (4). For example, the creation of the Human Proteome Organization's Human Brain Proteome Project foster the effective international exchange of brain related proteomic data (4, 31). Complex diseases are now rapidly investigated by novel high-throughput biochemical technologies to uncover disease activity, clinical markers, and drug targets. Such diagnostic technologies will lead to personalized medicine.

Opposite to the genome, the proteome is composed of an active array of molecules constantly being modified and with special localization. Proteomic approaches are able to characterize also post-translational modifications, by which the cell quickly modifies protein functions. Protein profiling and identification techniques using advanced mass spectrometry and bioinformatics can lead to the discovery, identification, and characterization of protein biomarkers (32). Comprehensive proteomic profiling is able to identify thousands of proteins from various clinical samples. Comparison of the proteomes of patient's and control sample may result in the identification of diseases specific proteins. Similar to gene expression profiling, several protein profiling techniques do not require *a priori* knowledge of candidate proteins (4). New tools for highly specific, sensitive, parallel protein analyses both in body fluids and tissue extracts will make a profound impact on clinical diagnostics in the near future. Profiling the proteomes of diseased and healthy tissues allows for the discovery of peptide or protein molecular change, which potentially reveals information on pathogenesis or diagnosis, or both (4).

Current proteomic research follows at least two pathways. In the first, the identification of proteins in patient's samples can be used as diagnostic or prognostic disease markers. The second goal is to discover cellular proteins related to the response to various therapies. Once tools for conducting comprehensive proteome analysis became available, much of the interest turned towards analyzing proteins for the purpose of finding novel biomarkers of diseases, such as cancer (33). In the future, the ability to routinely identify thousands of proteins in the body fluids will be available. Because of the unique protein content of these samples, strategies for removing highly abundant proteins needed to be developed (33). Several highly expressed proteins, particularly serum albumin, transferrin, and immunoglobulins, often mask lower abundance proteins. Thus purification measures to fractionate and better resolve protein population have to be undertaken. One of the biggest challenges in finding biomarkers in clinical samples was throughput. To obtain the identification of thousands of proteins in a sample a high throughput technique is needed (33). Unless there is a major breakthrough in technology, biomarker discovery studies using comprehensive proteomic identification will remain low-throughput compared to genomic microarray analysis.

Initial proteomic studies relied on 2D-gel electrophoresis, which separates proteins based on isoelectric point and molecular weight (4). This process has limited reproducibility, is complicated and not robust. Moreover weakly soluble proteins cannot be easily resolved and only a tiny portion of the proteome can be effectively stained. Another shortcoming of this method is that low-level expressed proteins can be masked by greater expression within a similar molecular weight or isoelectric point, or both. Edgar et al. applied this approach to the hippocampal proteome of schizophrenia patients (4, 34).

Proteomics advanced dramatically with the advent of mass spectrometric analysis for peptides (MALDI) (27, 35). There are four steps in mass spectrometry. First, the ion source generates ionized proteins from the sample. Second, the mass analyzer sorts and resolves proteins based on their mass/charge ratio. Third, the ion detector spots the ions and composes data on the ion mass/charge ratio, quantity, and time of flight (TOF), or the time it took to reach the detector. Finally, bioinformatic analysis allows interpretation of the raw data (30, 31, 32). After a mass spectrometry run, in a process called peptide mass fingerprinting, the peptides are arranged into several databases to allow protein identification. Although peptide mass fingerprinting is a method of protein identification, it often requires extensive and often complex purification, as it tends to interpret protein match by peptide masses rather than by sequences (30, 31, 32). Mass spectrometry has evolved to incorporate tandem mass spectrometric technology that permits effective sequencing. The MALDI-TOF is an advanced technology, a cutting-edge proteomic tool with direct amino acid sequencing and characterization capabilities (27, 30, 32). As mass spectroscopy continues to improve, it may replace immunoassays as the best method for measuring specific analytes in biologic samples.

Surface Enhanced Laser/Desorption Ionization (SELDI), a variation of MALDI, is a new generation of mass spectrometric analysis, and offers better accuracy with built in chromatography (27, 30 31, 35). The central technology platform is a protein chip mass spectrometer, which uses a powerful new approach (surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, or SELDI-TOF MS) to the analysis of complex protein mixtures such as serum and tissue extracts by profiling hundreds of proteins in the same time, thus creating characteristic patterns related to various cellular states or disease conditions (27, 35). Each chip contains a unique chromatographic surface for selective protein capture. The recent emergence of methods for rapid profiling thousands of protein markers by use of mass spectrometry has raised hope for the rapid identification of novel cancer biomarkers (27, 35). Specific questions concerns the reproducibility of SELDI-TOF, possible changes in protocols, calibration, and the ability of SELDI to detect low-abundance tumor markers. Additionally, observations that spectra can vary depending on analytical factors such as the time of processing have been noted during large profiling experiments (27, 36). Semmes et al demonstrated that relevant part of SELDI-TOF profiles can be measured reproducibly and used to distinguish a reference set of prostate cancer samples from controls (27, 37). This study was an encouraging step toward defining the analytical reproducibility of serum proteome profiling, although it highlighted the need for rigorous calibration of instruments and adherence to standardized technical procedures (27, 37). Although promising, the variation attributable to differences in sample collection and other sources of preanalytical bias that can be expected in routine clinical practice (27). A more comprehensive list of issues that must be addressed to understand the effects of preanalytical, analytical, and postanalytical factors on SELDI-TOF and matrix-assisted laser desorption/ionization (MALDI)-TOF profiles has recently been proposed (27, 38). The data provide evidence that preanalytical and analytical variation can affect profiled markers, and this must raise awareness of the strong risk for bias in serum profiling experiments that are not carefully controlled (27, 38).

Functional study employed the profiling and sequencing properties of tandem mass spectrometric analysis. However, several studies have demonstrated the potential of this technology in determining the complexities of the dynamic proteome (4). Sequence analysis can detect important post-translation protein modifications such as methylation, acetylation, sulfation, phosphorylation, ubiquitylation and glycosylation (4, 39, 40). Protein detection can be performed on microarrays, however heterogeneity and relative instability of proteins is a challenge (29, 41). Current research focus both on protein microarray construction and molecular strategies for specific and sensitive detection (29, 41). Antibody arrays can be used for protein expression studies and as diagnostic and discovery tools in autoimmunity (42).

One of the important tools in clinical proteomics are tissue microarrays (43). They allow to analyze hundreds of tissue specimens in the same time. Tissue arrays investigate the distribution of proteins directly at the disease site (43). The obtained results can be assessed manually or automatically and can be analysed together with clinical data (43).

An important issue in the search for potential diagnostic marker is validation of diagnostic method (44). At the validation stage, thousands of samples must be analyzed, therefore the throughput of the method is of tremendous importance. Another very important distinction between analysis of proteins at the discovery and validation stage involves the stringency in quantitation of the analytes (44, 45). In the discovery phase, measurements quantitating differences in analytes between samples usually are not particularly accurate or precise (44). The confidence level in very small differences in abundance between analytes in different samples is generally low for mass spectrometry measurements of complex peptide mixtures (44, 46). This low confidence level is tolerated for the sake of throughput and knowledge that any potential marker will necessarily go through a rigorous validation before it is ever used in a clinical setting (44). Another important issue in validation is low specificity of majority of the potential disease-specific biomarkers. Many of these proteins belong to acute-phase proteins whose concentrations universally change in response to infection or inflammation (44). Other proteins that have been reported to be potential biomarkers are

induced by factors such as diet and medication and may have absolutely no relationship to the underlying disease (44). Unfortunately these proteins generally rank among the highest abundant proteins in serum or plasma (44).

Correct patient recruitment is another challenge in biomarker discovery and validation (44). The challenge is considerably greater in validation simply because a much larger number of samples need to be analyzed. For validation study patients must be randomly selected with the correct demographics, medical history and lifestyle. Beyond patient samples, proper controls must also be acquired from individuals with similar demographics as the patients but being disease-free (44). Calculations need to be performed to determine the number of cases and controls that provide adequate statistical power once the results are analyzed. By incorporating proteins as standards, more robust absolute quantitation can be achieved by introducing the standards into sample preparation as early as possible, and even more important, to easily monitor multiple peptides derived from the same protein. To carry out global absolute quantitative proteomics, Silva *et al.* developed a method based on the discovery that the average mass spectrometry signal for the three most intense tryptic peptides per mole of any protein is constant within a coefficient of variation of less than $\pm 10\%$ (44, 47). Based on this hypothesis absolute protein amounts for all identified proteins in a sample can be calculated on the basis of an internal standard protein at known concentration. (44). All of the above practical issues can be important for the successful biomarker discovery and validation (44).

The number of diagnostic tools has been steadily expanding since the advent of modern medicine. New -omics technologies will allow thousands of results per sample to be generated. Novel clinical tests could be used to determine therapy type, duration, and dose; and its efficiency. Genomics and proteomics may also reveal distinct etiologies and subtypes for better classification of the disease. Applying genomic and proteomic methods to clinically accessible body fluids (e.g., serum, cerebrospinal fluid, and urine) would place valuable objective analytical power in the hands of the clinician (33, 41, 48). With personalized medicine, therapy will be based on individual patient characteristics that become known through bioinformatics. The expected results will give the response rates close to 100%, as well as increased survival rates, improved quality of life, cost savings, and reduced morbidity and mortality. For several years proteomics research has been expected to lead to the finding of new markers that will translate into clinical tests applicable to numerous clinical samples such as serum, plasma and urine (33, 41, 48). Attempts to implement technologies applied in proteomics, in particular protein arrays and surface-enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI-TOF MS), as diagnostic tools have initiated constructive discussions on opportunities and challenges. Genome and proteome-based research offer the promise of more effective diagnostic tools, greater understanding of an individual's healthcare needs, and targeted treatments for diseases that affect a vast majority of the population such as cancer, diabetes and cardiovascular disease (49, 50). The role of genomics and proteomics in health care is increasingly driven by the need for integrated approaches to disease prevention, earlier diagnosis and overall response to therapy – the new era of personalized medicine (24, 25, 50).

References

1. Vostrý M. Multiplex immunoassays chips and beads. *eJIFCC* 2009 Vol 20 No 4
2. Rados C. Genomics and medical devices: a new paradigm for health care. FDA Consumer. http://findarticles.com/p/articles/mi_m1370/is_6_39/ai_n27869192/
3. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860–921. doi:10.1038/35057062
4. Lakhan SE. Schizophrenia proteomics: biomarkers on the path to laboratory medicine? *Diagnostic Pathology* 2006, 1:11doi:10.1186/1746-1596-1-11
5. Homfray T, Holland T, Patton M. A new X-linked mental retardation syndrome. *Clin Dysmorphol* 1995; 4: 289-293
6. Rosenson RS. New technologies personalize diagnostics and therapeutics. *Curr Atheroscler Rep.* 2010;12:184-6
7. Bentley DR. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* 2006; 16: 545–552
8. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet.* 2010; 11: 31-46
9. Feng X, Du W, Luo Q, Liu BF. Microfluidic chip: next-generation platform for systems biology. *Anal Chim Acta.* 2009; 650: 83-97

10. Ho YP, Chen HH, Leong KW, Wang TH. Combining QD-FRET and microfluidics to monitor DNA nanocomplex self-assembly in real-time. *J Vis Exp*. 2009; 30 pii: 1432. doi: 10.3791/1432
11. Sims AH. Bioinformatics and breast cancer: what can high-throughput genomic approaches actually tell us? *J Clin Pathol*. 2009; 62: 879-85
12. Ferreira EN, Maschietto M, Silva SD, Brentani H, Carraro DM. Evaluation of quantitative rt-PCR using nonamplified and amplified RNA. *Diagn Mol Pathol*. 2010; 19: 45-53
13. Redon R, Carter NP. Comparative genomic hybridization: microarray design and data interpretation. *Methods Mol Biol*. 2009; 529:37-49
14. O'Connor L, Glynn B. Recent advances in the development of nucleic acid diagnostics. *Expert Rev Med Devices*. 2010; 7:529-39
15. Singhal S, Miller D, Ramalingam S, Sun SY. Gene expression profiling of non-small cell lung cancer. *Lung Cancer*. 2008; 60:313-24
16. MacPhee WN, Current IA. Progress in pharmacogenetics and individualized immunosuppressive drug dosing in organ transplantation. *Curr Opin Mol Ther*. 2010; 12: 270-83
17. Rubinsztein DC. Functional genomics approaches to neurodegenerative diseases. *Mamm Genome* 2008; 19: 587-90
18. Eshaque B, Dixon B. Technology platforms for molecular diagnosis of cystic fibrosis. *Biotechnol Adv*. 2006; 24: 86-93
19. Ferrer M, Sanz ML, Sastre J, Bartra J, del Cuvillo A, Montoro J, Jáuregui I, Dávila I, Mullol J, Valero A. Molecular diagnosis in allergology: application of the microarray technique. *J Investig Allergol Clin Immunol*. 2009;19 Suppl 1:19-24
20. Diosdado B, van Oort E, Wijmenga C. Coelionomics: towards understanding the molecular pathology of coeliac disease. *Clin Chem Lab Med*. 2005; 43: 685-95
21. Zheng J, Moorhead M, Weng L, Siddiqui F, Carlton VE, Ireland JS, Lee L, Peterson J, Wilkins J, Lin S, Kan Z, Seshagiri S, Davis RW, Faham M. High-throughput, high-accuracy array-based resequencing. *Proc Natl Acad Sci U S A*. 2009; 21: 6712-7
22. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, et al. NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Res*. 2009; 37: D885–D890.
23. Kiefer J, Yin HH, Que QQ, Mousset S. High-throughput siRNA screening as a method of perturbation of biological systems and identification of targeted pathways coupled with compound screening. *Methods Mol Biol*. 2009; 563: 275-87
24. Cowin PA, Anglesio M, Etemadmoghadam D, Bowtell DD. Profiling the cancer genome. *Annu Rev Genomics Hum Genet*. 2010 ;11: 133-59.
25. Auffray C, Chen Z, Hood L. Systems medicine: the future of medical genomics and healthcare. *Genome Med*. 2009; 1: 2-9
26. Molloy RM, Mc Connell RI, Lamont JV, FitzGerald SP. Automation of biochip array technology for quality results. *Clin Chem Lab Med*. 2005; 43(12):1303-13.
27. Stephen R. Master. Diagnostic Proteomics: Back to Basics? *Clinical Chemistry*. 2005; 51: 1333-1334

28. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 2001; 292: 929-934
29. Stoevesandt O, Taussig MJ, He M. Protein microarrays: high-throughput tools for proteomics. *Expert Rev Proteomics*. 2009; 6: 145-57
30. Ocak S, Chaurand P, Massion PP. Mass spectrometry-based proteomic profiling of lung cancer. *Proc Am Thorac Soc*. 2009; 6:159-70
31. Hamacher M, Meyer HE. HUPO Brain Proteome Project: aims and needs in proteomics. *Expert Rev Proteomics* 2005; 2: 1-3
32. Walther TC, Mann M. Mass spectrometry-based proteomics in cell biology. *J Cell Biol*. 2010; 190: 491-500
33. Issaq HJ, Xiao Z, Veenstra TD. Serum and plasma proteomics. *Chem Rev*. 2007; 107: 3601-20
34. Edgar PF, Schonberger SJ, Dean B, Faull RL, Kydd R, Cooper GJ. A comparative proteome analysis of hippocampal tissue from schizophrenic and Alzheimer's disease individuals. *Mol Psychiatry* 1999; 4: 173-178
35. Conrads TP, Zhou M, Petricoin EF, 3rd, Liotta L, Veenstra TD. Cancer diagnosis using proteomic patterns. *Expert Rev Mol Diagn*. 2003; 3:411-420
36. Coombes KR, Morris JS, Hu J, Edmonson SR, Baggerly KA. Serum proteomics profiling—a young technology begins to mature. *Nat Biotechnol* 2005; 23:291-292
37. Semmes OJ, Feng Z, Adam BL, Banez LL, Bigbee WL, Campos D, et al. Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. *Clin Chem* 2005; 51:102-112
38. Hortin GL. Can mass spectrometric protein profiling meet desired standards of clinical laboratory practice? *Clin Chem* 2005; 51: 3-5
39. Medzihradszky KF: Peptide sequence analysis. *Methods Enzymol* 2005; 402:209-244
40. Wang D, Kalume D, Pickart C, Pandey A, Cotter RJ Identification of Protein Ubiquitylation by Electrospray Ionization Tandem Mass Spectrometric Analysis of Sulfonated Tryptic Peptides. *Anal Chem* 2006; 78: 3681-3687
41. Rimini R, Schwenk JM, Sundberg M, Sjöberg R, Klevebring D, Gry M, Uhlén M, Nilsson P. Validation of serum protein profiles by a dual antibody array approach. *J Proteomics* 2009; 73: 252-66
42. Plebani M, Pittoni M, Celadin M, Bernardi D, Mion MM. Recent advances in diagnostic technologies for autoimmune diseases. *Autoimmun Rev*. 2009; 8: 238-43
43. Kononen J, Bubendorf L, Kallioniemi A, Bärnlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med*. 1998; 4: 844-7
44. Ye X, Blonder J, and Veenstra TD. Targeted proteomics for validation of biomarkers in clinical samples. *Brief Funct Genomic Proteomic*. 2009; 8: 126–135
45. Kaiser J. Proteomics. Public-private group maps out initiatives. *Science*. 2002; 3: 827
46. Johann DJ Jr, Blonder J. Biomarker discovery: tissues versus fluids versus both. *Expert Rev Mol Diagn*. 2007; 7: 473-5

47. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics*. 2006; 5: 144-56
48. Pejcic M, Stojnev S, Stefanovic V. Urinary proteomics--a tool for biomarker discovery. *Ren Fail*. 2010; 32: 259-68
49. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med*. 2010; 363:166-76
50. Calzone K, Wattendorf D, Dunn BK. The application of genetics and genomics to cancer prevention. *Semin Oncol*. 2010; 37: 407-18