Proteomics: A new perspective for cancer

Basavaradhya Sahukar Shruthi, Palani Vinodhkumar[1], Selvamani[2]

Department of Oral and Maxillofacial Pathology and Microbiology, Vishnu Dental College, Vishnupur, Bhimavarm, Andhra Pradesh, [1]Department of Pedodontics, Sri Balaji Dental College and Hospital, Chennai, Tamil Nadu, [2]Department of Oral and Maxillofacial Pathology and Microbiology, College of Dental Sciences, Davangere, Karnataka, India

Abstract

In the past decades, several ground breaking discoveries in life science were made. The completion of sequencing the human genome certainly belongs to the key tasks successfully completed, representing a true milestone in the biomedicine. The accomplishment of the complete genome also brings along a new, even more challenging task for scientists: The characterization of the human proteome. Proteomics, the main tool for proteome research, is a relatively new and extremely dynamically evolving branch of science, focused on the evaluation of gene expression at proteome level. Due to the specific properties of proteins, current proteomics deals with different issues, such as protein identification, quantification, characterization of post-translational modification, structure and function elucidation, and description of possible interactions. This field incorporates technologies that can be applied to serum and tissue in order to extract important biological information in the form of biomarkers to aid clinicians and scientists in understanding the dynamic biology of their system of interest, such as a patient with cancer. The present review article provides a detail description of proteomics and its role in cancer research.

Key Words: Biomarkers, electrophoresis, neoplasia, spectrometry, translation

INTRODUCTION

Proteins are the building blocks of the living organism made of different amino acids. They vary greatly in their function, stability, three dimensional structures, and are considered as the final gene products of most of the genes. The complete genome sequences of many organisms including higher organisms like human beings are known today. However, the functions of most of the genes remain obscure. The availability of this novel unexplored ocean of data has led some scientists to designate the 21st century as the post-genomic or proteomic era.

The term proteome indicates the total set of proteins encoded by a genome. Proteomics involves the surveying of global protein composition of a cell or organism. It is necessary to monitor the level along with the activity of proteins.[1] Proteomic data are tremendously useful in classifying cells and tissues in disease states and understanding different biological mechanisms. The structure, interactions, and functions of all proteins within cells and organisms can be identified by utilizing methods of protein measurement. This field incorporates technologies

Access this article online	
Quick Response Code:	**Website:** www.advbiores.net
	DOI: 10.4103/2277-9175.180636

that can be applied to serum and tissue in order to extract an important biological information in the form of biomarkers to aid clinicians and scientists in understanding the dynamic biology of their system of interest, such as a patient with cancer.[2]

As an important biological indicator of cancer status and progression for the physiological state of the cell at a specific time, biomarkers represent powerful tools for monitoring the course of cancer and gauging the efficacy and safety of novel therapeutic agents.[3] Thus the better understanding of an emerging field of proteomics will provide the knowledge needed for identification of these biomarkers, targeting specific protein pathways and in turn improve the health-care.[1]

DEFINITION AND GOALS OF PROTEOMICS

According to Kiernan "Proteomics is the use of quantitative protein-level measurements of gene expression to characterize biological processes (e.g., Disease processes and drug effects) and decipher the mechanisms of gene expression control."[4]

The end result of genome sequencing project suggest that the human genome contains more than 30,000 unique genes, which could give rise to as many as 100,000 distinct protein products. These staggering numbers have shifted the focus of new biotechnologies from those that facilitate the characterization of single genes or proteins, to methods that enable the rapid monitoring of all possible proteins contained in a living organism.[5]

Proteomics is not only concerned with all the proteins in any given cell, but also the set of all protein isoforms and modifications, the interactions between them, the structural description of proteins and their higher-order complexes, and most of the "post-genomic" data.[6]

The field of proteomics emerged with the goals of developing and applying the methods for the global analysis of protein expression and function.[7] Current goals of proteomic research are more varied and directed towards the systematic determination of diverse properties of proteins in various physiological and pathological conditions.[8]

It is anticipated that the creation of effective methodologies for the rapid and parallel analysis of proteins will accelerate the functionalization of biomolecules and thus, enabling the discovery of new biomarkers and therapeutic targets for the diagnosis and treatment of diseases of human and generally increase our mechanistic understanding of biological processes.[7]

NEED FOR PROTEOMICS

Functional genomics has been used to describe the analysis of changes in gene expression in response to various experimental conditions.[9] Comprehensively, it is whether and when a particular gene is transcribed and translated, at what rate, under which specific circumstance, and its functional end result, which has remained undetermined from studies at the level of genomic sequence.[10] This type of analyses have played an important role in studying physiology and pathophysiology of various diseases.[5] Attention must be focused on gene expression and functions of the proteins they encode, once the initial stage of genome sequencing and gene discovery is completed.[10] However, despite great advances in bioinformatics, it is still difficult to predict genes accurately from genomic data.[9] Some of the limitations of genomic analysis are that they cannot provide complete information of cellular, subcellular, and intercellular functions, in which proteins, not genes govern the functions,[8] and there is no strict linear relationship between genes and the protein complement or "proteome" of a cell.[11] In addition, many genes are pseudogenes that are no longer expressed in cells.[9]

It is now generally accepted that analyses undertaken at the level of proteins are necessary for the following reasons.[1,10]

- There is poor correlation between mRNA (messenger Ribonucleic acid) abundance and corresponding protein levels. This implies that protein levels cannot simply be predicted from corresponding mRNA levels
- Some mRNA molecules are non-coding and do not give rise to any protein products
- Some primary mRNA transcripts undergo alternative splicing; therefore, the gene may give rise to multiple protein products
- Virtually all eukaryotic proteins undergo post-translational modifications. These modifications, which potentially have enormous functional consequences, cannot always be predicted from gene sequences
- The translocation of protein from its site of synthesis to the site of activity cannot always be deduced from sequence data
- Proteins themselves may be degraded and vary greatly in stability
- Protein function cannot always be reliably predicted from sequence information.

METHODS OF PROTEIN MEASUREMENT AND BIOMARKER IDENTIFICATION

Various technologies have been developed to identify proteomic measurement [Figure 1].[12] Proteomic
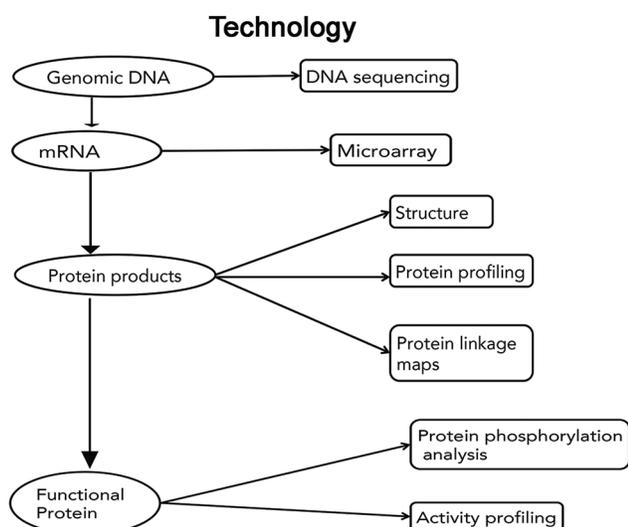
Technology

Figure 1: Review of technologies used for identification of proteins

data are tremendously useful in classifying cells and tissues in disease states and understanding different biological mechanisms.[12,13] Scientists identify the structure, interactions, and functions of all proteins within the cells and organisms by utilizing methods of the protein measurement. The current focus in proteomics is to define the nature of proteins and their levels of expression in cells using the protein detection technologies. Clinical samples of proteins can be analyzed through procedures, such as mass spectrometry, 2-Dimensional polyacrylamide gel electrophoresis (2D PAGE), and protein arrays in order to help compare the protein variability between samples.[14]

2D page

The most commonly used technology for analyzing protein variability between samples is the 2D PAGE method. Samples are denatured, processed and separated according to their isoelectric points in the first dimension. This mixture is then separated in another gel matrix, i.e., second dimension according to the molecular weight. Individual protein spots are then stained. The detected proteins are cut and digested into fragments. These fragments are analyzed with high-resolution mass spectrometry. The digested fragments are then matched from known protein databases.[15]

Mass spectrometry

The most accurate and reliable method for analyzing a biological sample from a patient is called mass spectrometry. This method identifies proteins in a substance by their mass and charge.[16,17] Matrix-assisted laser desorption and ionization with time-of-flight detection mass spectrometry (MALDI-TOF) and surface-enhanced

laser desorption and ionization with time-of-flight spectrometry (SELDI-TOF) are two of the methods currently being employed. MALDI techniques immobilize protein samples in an energy absorbing matrix (chemical) on a chip or plate. The entire repertoire of proteins in the sample interacts with the matrix from which a selected subset of proteins is bound to, a function of the composition of the selected matrix. SELDI technology uses selective surfaces for binding a subset of proteins based on absorption, partition, electrostatic interaction or affinity chromatography on a solid-phase protein chip surface.[2] Proteomic techniques utilizing mass spectrometry include: (i) Protein-protein interactions; (ii) post-transitional modifications; (iii) structural proteomics; (iv) protein quantitation or differential modifications; and (v) protein identification. These techniques are used by laboratories to isolate proteins and identify protein characteristics to help clinicians determine between the healthy and disease states of the patient.[18]

Protein arrays

A protein microarray is a piece of nitrocellulose coated glass slide on which different molecules of protein have been bound at separate locations.[19] Protein arrays use antibodies of known affinity and specification, and they are stamped on the surface of aptamers. This method allows the observation of the biochemical activities of thousands of proteins.[20] Prior knowledge of molecules is needed in order to use this method for discovering biomarkers. One aspect of the protein array approach that distinguishes it from DNA microarrays is its ability to identify distinct protein isoforms that might be critical in detecting disease pathogenesis.[14] Discovering these specific biomarkers is clinically useful in indicating alterations between a normal and diseased sample. This information is helpful to clinicians in the diagnosis of diseases, as well as in the monitoring of patient's responses to therapy.[21]

Protein bioinformatics

Experiments carried out in a real laboratory need to be complemented by virtual experiments done on the computer. In addition to the software packages for analyzing the electrophoretic separation, bioinformatic tools have been developed. Some of these are available via the internet with links to many provided from the Expert Protein Analysis System (ExPASy) proteomics server.[21]

This software allows not only the identification of proteins but further characterization ranging from the calculation of basic physicochemical properties to the prediction of potential post-translational modifications and three-dimensional structures.[21]

Biomedical applications

Despite tremendous advances in our understanding of the molecular basis of various diseases, substantial gaps remain both in our understanding of disease pathogenesis and in the development of effective strategies for early diagnosis and treatment. The current interest in proteomics is due to the prospects that proteomic approach to any disease will overcome some of the limitations of other approaches.[22] The field of proteomics studies proteins in an effort to catalog them and to understand their role in biology and pathology so they may be applied to early diagnosis and to optimizing treatments.[2]

Ever since, proteomics was proven to be capable of characterizing a large number of differences in both protein quality and quantity, it has been applied in various areas of biomedicine, ranging from the deciphering molecular pathogenesis of diseases to the characterization of novel drug targets and the discovery of potential diagnostic biomarkers.[23]

PROTEOMICS AND CANCER

Cancer is a multifaceted disease which results from dysregulated normal cellular signaling networks that control cell behaviors, such as proliferation and apoptosis, caused by genetic, genomic and epigenetic alterations at the cell or tissue levels.[24]

More than 11 million people are diagnosed with cancer every year. It is estimated that there will be 16 million new cases every year by 2020. From a total of 58 million deaths worldwide in 2005, cancer accounts for 7.6 million (or 13%) of the global mortality. Deaths from cancer in the world are projected to continue rising, with an estimated 9 million people dying from cancer in 2015 and 11.4 million dying in 2030.[3]

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.[25] Prevention, early detection and early intervention are the primary aims of oncologists and cancer biologists.[2] If genes are viewed as the master controllers of cellular behavior, proteins are the effectors, and as such, protein expression and activity must comprise the molecular basis of health or disease. Specifically related to cancer, expressed proteins direct tumor growth, invasion, metastases, interaction with surrounding cells, and response to therapy.[26] Uncovering the protein signaling network changes, including cell cycle gene network in cancer, aids in understanding the molecular mechanism of carcinogenesis, cancer progression and metastasis and thus identifies the characteristic signaling network

signatures unique for different cancers and specific cancer subtypes. Signaling network alterations accumulate at each stage of carcinogenesis that results from genetic, epigenetic and environmental changes and is viewed as a multi-step model of carcinogenesis.[24]

Oncoproteomics is a branch of proteomics which includes the study of proteins and their interactions in a cancer cell by proteomic technologies. There is an intense interest in applying proteomics to foster an improved understanding of cancer pathogenesis, develop new tumor biomarkers for diagnosis, and early detection using proteomic portrait of samples. Oncoproteomics has the potential to revolutionize clinical practice, including cancer diagnosis and screening based on proteomic platforms as a complement to histopathology, individualized selection of therapeutic combinations that target the entire cancer-specific protein network, real-time assessment of therapeutic efficacy and toxicity, and rational modulation of therapy based on changes in the cancer protein network associated with prognosis and drug resistance.[3]

Most currently available screening tests for cancers lack high sensitivity and specificity to be useful in screening the general population, so the differentiation between some benign and malignant tumors is still a clinical challenge. The advent of oncoproteomics has provided the hope of discovering novel biomarkers for use in screening, early diagnosis, and prediction of response to therapy. Like normal cells, most cancer cells use multiple redundant intracellular signaling pathways to ensure the maintenance and viability of functions critical to their survival. Thus, cellular pathways that are integral to cell function, survival, proliferation, and receptor expression are potential targets for therapeutic intervention. Clinicians might recommend combinations of molecularly targeted agents and other therapies on the basis of an individual patient's proteomic profile.[3]

EARLY DIAGNOSIS OF CANCER

Malignant transformation involves alterations in protein expression with subsequent clonal proliferation of the altered cells. These alterations can be monitored at the protein level, both qualitatively and quantitatively. Protein signatures in cancer provide valuable information that may be an aid to more effective diagnosis, prognosis, and response to therapy.[25] The major challenge of proteomic research is the limited assay sensitivity of analyzing cell proteins. Furthermore, the proteins involved in cellular homeostasis, metabolism and structure are abundant and are present 10,000-100,000 fold greater than proteins involved in signaling networks

in an individual cell. Therefore, detection and quantification of these cell signaling proteins poses a great challenge.[24]

The ability to characterize proteins within complex biological fluids such as serum, plasma, nipple aspirate fluid and urine by proteomic technologies has reached a point where hundreds of species can be identified in a rapid fashion, which brings the greater possibility of identifying the desired biomarkers of cancer.[3,27]

Amongst the different fluids of the body, the human plasma is not only the primary clinical specimen but also represents the largest and deepest version of the human proteome present in any sample. The human plasma proteome holds the promise of a revolution in diagnosis of cancer and therapeutic monitoring, provided that major challenges in proteomics like early detection, prognosis of disease and related disciplines can be addressed.[28]

Although studies concerned with the identification of novel antigens or markers for diagnostic, prognostic, or therapeutic use have been paramount, molecules and processes implicated in carcinogenesis *per se* are increasingly being investigated. Most tumor markers in current use were identified from protein-based approaches, from the identification in the 1800s of an abnormal urinary precipitate in multiple myeloma (Bence-Jones protein) to the generation of tumor-specific antibodies against epithelial cancer cell lines. Genetic markers detected cytogenetically or by mutation detection, are also now entering clinical practice, but some changes likely to be important in carcinogenesis, and its diagnosis such as abnormal expression of proto-oncogenes may not be associated with a detectable genetic lesion.[29]

Biomarkers were initially discovered using the conventional methods like protein distillation, Enzyme Linked Immuno Sorbent Assay, Western Blot and Gel Electrophoresis. As they were not very specific and sensitive methods, the newer and advanced methods were invented. The currently used analyzers in proteomics for biomarker discovery are 2D PAGE, mass spectrometry, MALDI, Electrospray Ionization, SELDI-TOF. The important phase of the biomarker discovery process is the validation phase, which requires a clinical assay to be developed and extensively tested on thousands of clinical samples.[23,24] Processing of obtained sample is illustrated in Figure 2.[25]

DETERMINING TUMOR AGGRESSIVENESS

Clinically, it is obvious that some tumors metastasize and/or progress relentlessly, despite all therapeutic
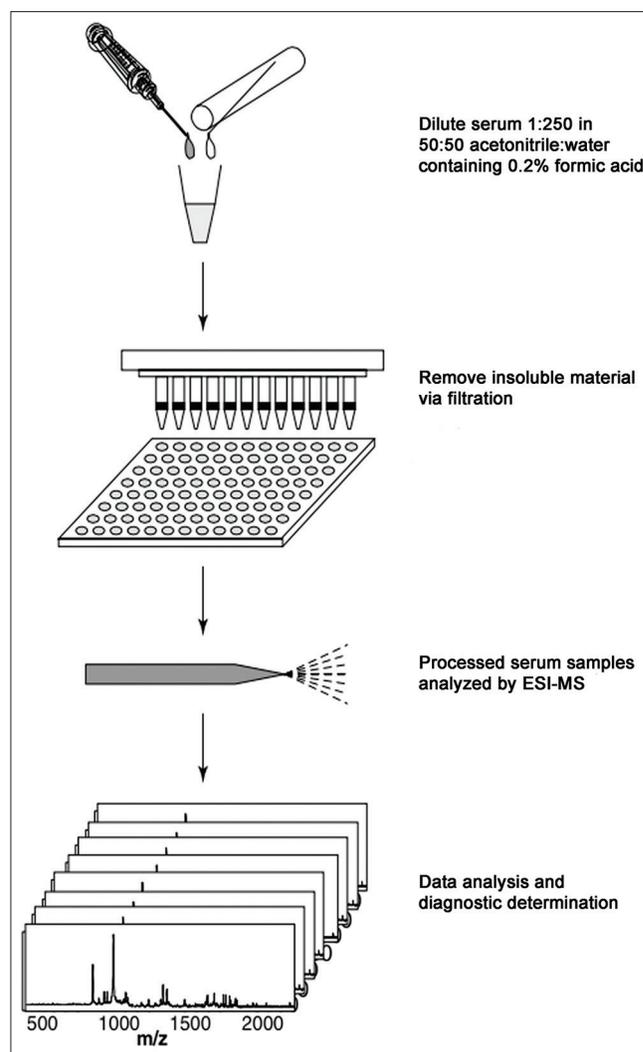
Figure 2: Processing of fluid sample (In this method, raw serum is diluted and filtered to remove particulates. The proteomic pattern of the filtered samples is then acquired using electrospray ionization mass spectrometry. As with surface-enhanced laser desorption and ionization with time-of-flight spectrometry-Mass spectrometry 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)

intervention, whereas other tumors grow slowly and never metastasize. Regardless of such heterogeneity in tumor behavior, cancer is treated as if it were a single disease. Because of this monolithic practice, many patients are unnecessarily exposed to aggressive therapy that can be associated with life-long morbidity or even mortality. Likewise, if patients with very aggressive disease could be identified, it would be appropriate to treat even more aggressively. Although proteomic technologies have not been used to segregate cancer based on clinical aggressiveness, identification of predictive peaks may lead to understanding of pathways that govern tumor aggressiveness.[26]

INDIVIDUALIZED THERAPY

The recent progress of proteomics has opened new avenues for cancer-related biomarker discovery. With the advent of new and improved proteomic technologies such as the development of quantitative proteomic methods, high resolution, high-speed, high-throughput, high-sensitivity mass spectrometry and protein chips, as well as advanced bioinformatics for data handling and interpretation, it is possible to discover biomarkers that are able to reliably and accurately predict outcomes during cancer treatment and management.[3]

Molecular markers used in various tumors to guide the therapy

Proteomic systems using tumor lysates or antibody arrays are being developed that can determine the activation of many or potentially all known kinases within specific tumors. Identification of kinases that are driving growth or aggressiveness of particular tumors could be used to target specific kinase inhibitors as therapy in these tumors. Mass spectrometry may identify markers predictive of tumor aggressiveness or metastatic potential. Better prediction of tumor behavior alone will allow better clinical decision-making with more appropriate treatment.[26]

CONCLUSION

Obstacles and future challenges

The field of proteomics has yielded a set of technologies and analytical techniques that are significantly advancing the field of cancer diagnostics. These technologies are found to be an efficient means of identifying new biomarkers for the early detection of cancer, and promise hope of new serological screening methods for diagnosis. Though proteomics is found to be complementing genomics-based approaches, providing additional information, it presents with various challenges in technical aspects, data collection and its inference. For example, there is no technique equivalent of polymerase chain reaction for amplification of low-abundance proteins, so a range of detection from one to several million molecules per cell is needed.[23]

Certain technological processes, particularly protein separation and analysis, are inherently skill-based and remain difficult to automate. Separation techniques such as capillary electrophoresis may be more amenable to automation but are unlikely to replace two-dimensional electrophoresis with its superior resolving power. Once the proteins are identified, bio informatics plays an important role in expanding the initial protein information, thus making it a crucial step, in which mishandling of data should be avoided to prevent the further mishaps.[29]

With the introduction of national proteomic funding initiatives, proteomics based approaches should be allowed to realize their potential in biomedical research and translation into clinical practice, thus making it to emerge as a powerful tool for combating the disease and maintaining the health.[29]

REFERENCES

1. Clark DP. Proteomics: The global analysis of proteins. Molecular Biology-Understanding the Genetic Revolution. Ch. 26. London: Elsevier Publication; 2005. p. 718-25.
2. Posadas EM, Simpkins F, Liotta LA, MacDonald C, Kohn EC. Proteomic analysis for the early detection and rational treatment of cancer: Realistic hope? Ann Oncol 2005;16:16-22.
3. Cho WC. Contribution of oncoproteomics to cancer biomarker discovery. Mol Cancer 2007;6:25.
4. Kiernan UA. Biomarker rediscovery in diagnostics. Expert Opin Med Diagn 2008;2:1391-400.
5. Patricelli MP. Activity-based probes for functional proteomics. Brief Funct Genomic Proteomic 2002;1:151-8.
6. Tyers M, Mann M. From genomics to proteomics. Nature 2003;422:193-7.
7. Aebersold R, Cravatt BF. Proteomics: Advances, applications and the challenges that remain. Trends Biotechnol 2002;20:S1-2.
8. Patterson SD, Aebersold RH. Proteomics: The first decade and beyond. Nat Genet 2003;33:311-23.
9. Klein JB, Thongboonkerd V. Overview of proteomics. Contrib Nephrol 2004;141:1-10.
10. Pardanani A, Wieben ED, Spelsberg TC, Tefferi A. Primer on medical genomics. Part IV: Expression proteomics. Mayo Clin Proc 2002;77:1185-96.
11. Pandey A, Mann M. Proteomics to study genes and genomes. Nature 2000;405:837-46.
12. Pierce JD, Fakhari M, Works KV, Pierce JT, Clancy RL. Understanding proteomics. Nurs Health Sci 2007;9:54-60.
13. Kenyon GL, DeMarini DM, Fuchs E, Galas DJ, Kirsch JF, Leyh TS, *et al.* Defining the mandate of proteomics in the post-genomics era: Workshop report. Mol Cell Proteomics 2002;1:763-80.
14. Resing KA, Ahn NG. Proteomics strategies for protein identification. FEBS Lett 2005;579:885-9.
15. Mattow J, Schmidt F, Höhenwarter W, Siejak F, Schaible UE, Kaufmann SH. Protein identification and tracking in two-dimensional electrophoretic gels by minimal protein identifiers. Proteomics 2004;4:2927-41.
16. Gu S, Chen X. Precise proteomic identification using mass spectrometry coupled with stable isotope labeling. Analyst 2005;130:1225-31.
17. McCormack AL. Mass spectrometry in proteomics. Methods 2005;35:209-10.
18. Stults JT, Arnott D. Proteomics. Methods Enzymol 2005;402:245-89.
19. Marik J, Lam KS. Peptide and small-molecule microarrays. Methods Mol Biol 2005;310:217-26.
20. Bertone P, Snyder M. Advances in functional protein microarray technology. FEBS J 2005;272:5400-11.
21. Fung ET, Wright GL Jr, Dalmasso EA. Proteomic strategies for biomarker identification: Progress and challenges. Curr Opin Mol Ther 2000;2:643-50.
22. Hanash S. Disease proteomics. Nature 2003;422:226-32.
23. Tambor V, Fucíková A, Lenco J, Kacerovský M, Rehácek V, Stulík J, *et al.* Application of proteomics in biomarker discovery: A primer for the clinician. Physiol Res 2010;59:471-97.
24. Zhang DY, Ye F, Gao L, Liu X, Zhao X, Che Y, *et al.* Proteomics, pathway array and signaling network-based medicine in cancer. Cell Div 2009;4:20.
25. Veenstra TD, Prieto DA, Conrads TP. Proteomic patterns for early cancer detection. Drug Discov Today 2004;9:889-97.

26. Yarbrough WG, Slebos RJ, Liebler D. Proteomics: Clinical applications for head and neck squamous cell carcinoma. Head Neck 2006;28:549-58.
27. Chung CH, Levy S, Chaurand P, Carbone DP. Genomics and proteomics: Emerging technologies in clinical cancer research. Crit Rev Oncol Hematol 2007;61:1-25.
28. Anderson NL, Anderson NG. The human plasma proteome: History, character, and diagnostic prospects. Mol Cell Proteomics 2002;1:845-67.
29. Banks RE, Dunn MJ, Hochstrasser DF, Sanchez JC, Blackstock W, Pappin DJ, *et al*. Proteomics: New perspectives, new biomedical opportunities. Lancet 2000;356:1749-56.