Academic Sciences International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 5, Issue 3, 2013

Review Article

USE OF GENOMICS AND PROTEOMICS IN PHARMACEUTICAL DRUG DISCOVERY AND DEVELOPMENT: A REVIEW

SHARMA NEHA, HARIKUMAR S.L.*

Rayat & Bahra Institute of Pharmacy, Sahauran, Kharar, District Mohali, Punjab 140104, India, S. L. Harikumar Rayat & Bahra Institute of Pharmacy, Sahauran, District Mohali, Punjab - 140104, India. Email: slharikumar@gmail.com

Received: 18 Mar 2013, Revised and Accepted: 23 May 2013

ABSTRACT

One of the most pressing issues facing the pharmaceutical industry is the tremendous dropout rate of lead drug candidates. Genomics and proteomics are today well established in drug discovery and development, in combination with combinatorial chemistry and high-throughput screening, are helping to bring forward a matchless number of potential lead compounds. Over the last two decades, several new genomic technologies have been developed in hopes of addressing the issues of target identification and lead candidate optimization. Proteomics is a technology platform that is gaining widespread use in drug discovery and drug development programs. Defined as the protein complement of the genome, the proteome is a varied and dynamic repertoire of molecules that in many ways dictates the functional form that is taken by the genome. We focus in this article on recent progress and innovations utilizing "omics" technologies to identify and validate drug targets, discover disease biomarkers, and design more effective drugs.

Keywords: Genomics, Proteomics, Pharmacogenomics, Drug discovery and development.

INTRODUCTION

The process of drug discovery is quite complex, integrating many disciplines, including structural biology, metabolomics, proteomics, and computer science, just to name a few. The process of drug discovery involves the identification of candidates, synthesis, characterization, screening, and assays for therapeutic efficacy. The process of drug development prior to clinical trials begins, when a compound has shown its value in the above tests. The process is generally quite tedious and expensive. [1]

The history of drug development over the past century has been the accumulation of knowledge and techniques that provide a progressively more detailed understanding of both the target and the compound that could become a drug.

Targets are usually proteins, either those occurring within the human body of in outside agents such as viruses and other pathogens. The major difficulty faced by drug researchers is understanding the complex che mical pathways involved in the disease process in order to find the most appropriate intervention point, and then to discover or design a compound that modifies the chemical process at that point.[2]

Apart from advances in technology and understanding of biological systems, drug discovery is still a long process with low rate of new therapeutic discovery. Information on the human genome, its sequence and what it encodes has been described as a potential windfall for drug discovery, promising to virtually eliminate the bottleneck in therapeutic targets that has been one limiting factor on the rate of therapeutic discovery.

However, data shows that "new targets" as opposed to "established targets" are more prone to drug discovery project failure in general. This data provide evidence that some thinking underlying a pharmaceutical industry trend beginning at the turn of the twenty-first century and continuing today which finds more risk aversion in target selection among multi-national pharmaceutical companies.

The process by which a new drug is brought to market stage is referred to by a number of names – most commonly as the development chain or "pipeline", and consists of a number of distinct stages. There are various estimates of the cost of each stage of the pipeline. [3]

The drug discovery process includes various steps which are as follows:

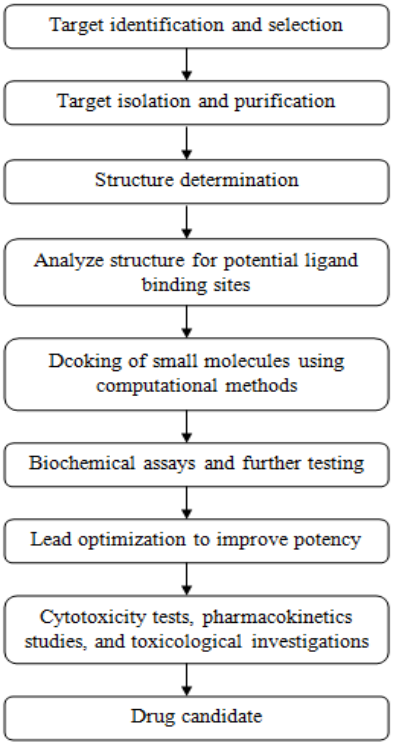

Fig. 1: Steps of Drug Discovery Process

Genomics is a science that studies the structure and function of genomes and, in particular, genes. A genome is an organism's complete genetic information. The genome is the collection of information that an organism can pass on to its offspring before birth.

Meanwhile, proteomics is essentially protein analysis and, until recently, could be described as a broad classification for a set of technology and bioinformatics platforms aimed at the comprehensive molecular description of the actual protein complement of a given sample. Presently, it is typically associated with systems biology. The considerable progress has been made in characterizing rapid posttranslational protein modifications in highly complex molecular signatures as key disease-related

biomarkers from experimental model systems or clinical samples. [4]

Difference between Genomics and Proteomics

Unlike the genome, which is relatively static, the proteome is not static; it changes constantly in response to tens of thousands of intra- and extracellular environmental signals. The proteome varies with various factors like health or disease, the nature of each tissue, the stage of cell development, and effects of drug treatments. As such, the proteome often is defined as "the proteins present in one sample (tissue, organism, and cell culture) at a certain point in time."

Proteomics runs parallel to genomics in many ways: Genomics starts with the gene and makes inferences about its products (proteins), whereas proteomics begins with the functionally modified protein and works back to the gene responsible for its production.

The sequencing of the human genome has increased interest in proteomics because while DNA sequence information provides a static snapshot of the various ways in which the cell might use its proteins, the life of the cell is a dynamic process. This new data increase the intrest of proteomics in the field of science, medicine, and most notably – pharmaceuticals.[5]

Role of Genomics in drug development

Drug development is defined in many pharmaceutical companies as the process of taking a new chemical lead through the stages necessary to allow it to be tested in human clinical trials. Molecular genetics reached human genetics about 1976, when the first human genes were cloned.[6]Transgenic methods, 'knock-outs' and 'knock-ins' began in about 1986, and in about 1996, database searching became a fruitful way to do genomic research.[7] The term'genomics' describes the discipline in genetics concerned with the study of the complete set of genes (genomes) of organisms.[8]The field includes efforts to determine the entire DNA sequence of organisms, fine-scale genetic mapping, studies of intragenomic phenomena such as heterosis, epistasis, pleiotropy and other interactions between loci and alleles within the genome. The definition of genomics does not include the investigation of the roles and a function of single genes is a primary focus of molecular biology. [9]

The primary goal of genomics research in the pharmaceutical industry in the 1990s was to identify not only new molecular targets but also more of them, and to be first to gain proprietary rights to use those targets.[10]Genomics explore new opportunities for the drug discovery, especially through technologies like high-throughput sequencing and characterization of expressed human genes.

Knowledge of all the human genes and their functions may allow effective preventive measures. The cause of common fatal diseases has been identified by genomics and it shows the potential to identify individuals who are particularly susceptible to a given disease long before that disease becomes apparent. It has positively impacted the drug research strategy and drug discovery development processes. The process has been made simpler and economical. Further innovations in this area are expected, which should take drug discovery research to a new level. [11]The figure 2 illustrate the potential contribution of genomics on drug development process.

Pharmacogenomics has its roots in pharmacogenetics. Whereas pharmacogenetics is the study of the linkage between an individual's genotype and that individual's ability to metabolize a foreign compound.

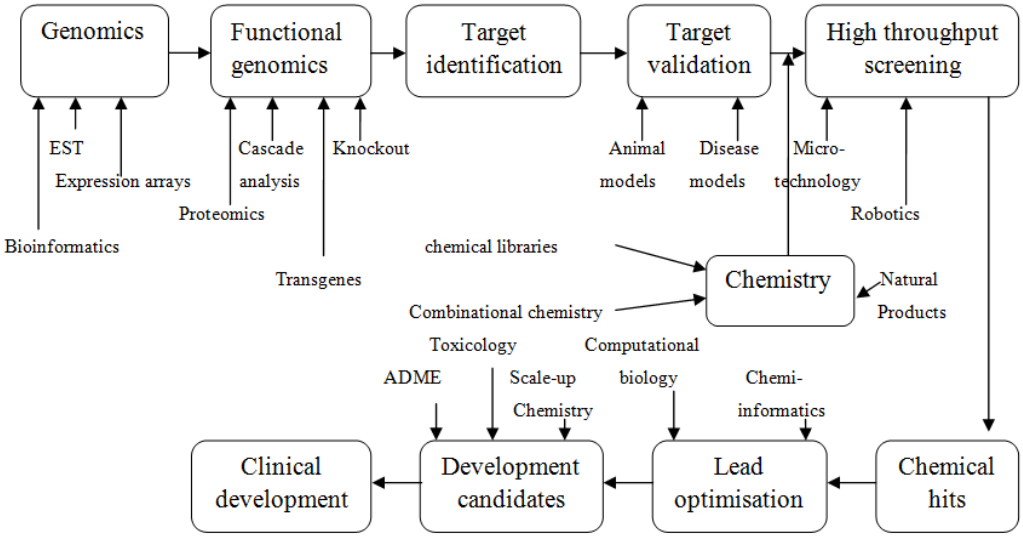

Fig. 2: Genomics to Clinical Drug Development

Pharmacogenomics is quite broad in scope, and is similar to molecular medicine, aiming to detect, monitor and treat the molecular causes of disease. Various genomics technologies such as gene sequencing, statistical genetics and gene expression analysis are used for the drugs in clinical development and trials. Since many diseases develop as a result of a network of genes failing to perform correctly, pharmacogenomics can identify the genes or loci which are involved in determining the responsiveness to a given drug. In this way, genetic characterization of patient populations is becoming an integral part of the drug discovery and development process. The main aim of genomics and Pharmacogenomics in clinical research and clinical medicine is that disease could be treated according to genetic and specific individual markers, selecting medications and dosages that are optimized for individual patients. [11]

Applying pharmacogenomics in the preclinical setting, one may start screening compounds with the least variation across individuals. When target gene is selected, the compound that works best overall against all its subtypes is chosen. Thus decreasing the uncertainties that patient stratification introduces at the FDA and in marketing, as well as the need for a genetic screening. Genomics may also be used to select out adverse effects before drugs enter the clinic. For example, the gene-expression pattern for the liver of an animal administered a drug can indicate whether gene pathways related to toxicity have been turned on. Variations in gene expression levels may prove just as useful as genetic variation in predicting drug response at any stage in the clinic and as a diagnostic. Pharmacogenetics data are vital during the development of a compound with a narrow therapeutic index or which is metabolized from a prodrug, as such information may influence decision of

whether to discontinue development or design trials to study clinical responses in individual polymorphic for the relevant enzyme.

Significant issues at the preclinical level usually need to be addressed. Problems of medicinal chemistry, developing drugs with the appropriate absorption, metabolism, distribution, and elimination profiles still have an empirical basis. Nonetheless, small molecule drugs directed toward targets discovered by genomics may soon account for a great majority of drugs introduced into the marketplace.

Pharmacogenomics benefits many stages of clinical drug development. It will significantly affect trial design, primarily through improved inclusion/exclusion criteria and more effective assessment of patient responses. During preclinical studies, the genes linked with drug metabolism could be genotyped in patients recruited for phase I trials. Any genotype that correlates with adverse effects could then be used to screen out relevant patients in subsequent trials. Furthermore, if efficacy data are collected during phase I trials, polymorphisms in the drug target gene could be typed in phase I participants to assess whether they are linked with side-effects or with variations in drug response. That analysis could obviously be further refined in phase II trials, enabling companies to undertake phase III trials in a subgroup of patients that responds well and exhibits fewer side-effects. The resultant drugs would be expected not only to have better efficacy, but also a better safety profile. At the clinical level, while the disease symptoms might appear to be uniform, individual-to-individual variations in these polygenic networks may make drugs healing for certain individuals while toxic for others.

The variation in gene can be correlated with differential responses to the same drug leads by pharmacogenomics, thereby hoping to accelerate novel drug discovery dramatically, by defining specific populations that will benefit from a drug. This approach may maximize the medical utility of existing pharmaceuticals and it could also rescue dead drugs. Several products that have failed in recent years in late stage clinical trials may on retrospective analysis be effective in subsets of patients, although at the time, there was no clear way of recognizing such subsets clinically. Consequently, traditional approaches that focus on broad groups of patients with a diagnosis (e.g. Alzheimer's disease) may need to be much more precisely divided into subsets of patients who may have a traditionally defined disease amenable to treatment based on a particular molecular target. These pharmacogenomic developments should lead to smaller, more rapid and cost-effective trials, and ultimately to more individually focused and effective therapeutics. [12]

Several genomics-based technologies already impact drug development by contributing to the identification of new targets, providing information for computer-aided design of lead compounds and target validation through disease models and predictive toxicology. Accordingly, “omics” technologies contribute to all stages of drug development, from target identification to target validation. Target identification is based upon molecular information derived from genome sequences and protein structures.

Overall, pharmacogenomic approaches offer interesting perspectives for molecular design and development of more specific drugs with significant benefits to patients. [13]

By applying genomics technology, companies can on average realize savings of nearly US$300 million and two years per drug, largely as a result of efficiency gains. Current research activities aim at going beyond the area of human genome sequencing to expand the list of identified proteins and genes. This, ultimately, is expected to help in improved understanding of disease mechanisms and the development of corresponding therapeutics.

Role of Proteomics in drug development

Recent advances in applied genomics helped in the target identification process, since it allowed for high throughput screening of expressed genes. However, studies have shown that there is a poor correlation between the regulation of transcripts and actual protein quantities. The reasons for this are that genome analysis does not account for post-translational processes such as protein modifications and protein degradation. Therefore, the methods employed in the drug-discovery process started to shift from genomics to proteomics. Proteomics is large-scale study of proteins, particularly their structures and functions.[14,15]Proteins are vital parts of living organisms, as they are the main components of the physiological metabolic pathways of cells.

The term "proteomics" was first coined in 1997to make an analogy with genomics, the study of the genes.[16]The word "proteome" is a blend of "protein" and "genome", and was coined by Marc Wilkins in 1994 while working on the concept as a PhD student.[17,18]Proteomics is a technology platform that is gaining widespread use in drug discovery and drug development programs. It provides effective means to identify biomarkers that have the potential to improve decision making surrounding drug efficacy and safety issues based on data derived from the study of key tissues and the discovery and appropriate utilization of biomarkers.[19]

With the accumulation of vast amounts of DNA sequences in databases, researchers are realizing that merely having complete sequences of genomes is not sufficient to elucidate biological function. A cell is normally dependent upon a multitude of metabolic and regulatory pathways for its survival. There is no strict linear relationship between genes and the protein complement or ‘proteome’ of a cell. Proteomics is complementary to genomics because it focuses on the gene products, which are the active agents in cells. For this reason, proteomics directly contributes to drug development as almost all drugs are directed against proteins. [20]

Table 1: Genomics and proteomics are synergetic [20]

Type of Information	Biological Carrier	Field of Application
Gene Sequence Data		
⇩	DNA	
Gene Expression Data		GENOMICS
⇩	RNA	
Gene Product Data		PROTEOMICS
⇩	PROTEIN	
Protein Function Data		

Proteins are the functional output of the cell and therefore might be expected to provide the most relevant information. The expression or function of proteins is modulated at many points from transcription to post-translation, which generally cannot be predicted from analysis of nucleic acids alone. There is poor correlation between the abundance of mRNA[21] transcribed from the DNA and the respective proteins translated from that mRNA and the transcript can be spliced in various ways to yield different protein forms. Extensive changes can also be introduced during or after translation—for example, the addition of specific carbohydrate side-chains or phosphorylation— leading to multiple protein products from a single gene.

Preliminary studies suggest an average number of protein forms per gene of one to two in bacteria, three in yeast, and three to more than six in human beings.[22] Thus, the human body may contain more than half a million modified proteins. The way in which gene and protein expression can be regulated or modified from transcription to post translation is expressed in figure 3.

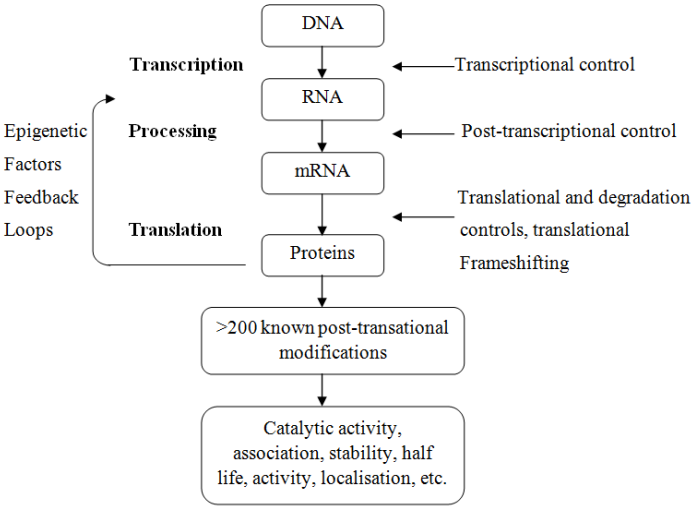

Fig. 3

Drug development is generally based around the desire to upregulate or downregulate a specific activity implicated in disease pathogenesis or in treatment-associated side effects. Most drugs exert their effects on proteins. The strategy of working forward from the gene has been used: a specific genetic lesion is identified and the resultant changes in protein structure, function, or expression are elucidated, so that a drug to counteract or correct such aberrations can be rationally designed. An example is the development of an inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukaemia. [23]

Proteomics has gained much attention as a drug development platform because disease processes and treatments are often manifesting at the protein level. Drugs ought to produce protein expression effects and, hence, the pattern of protein changes after drug application will give information about the mechanism of action, either for therapeutic or toxicological effects. Various drugs might be compared and grouped according to their signal in metabolistic pathways. Side effects can also be described if additional proteins are involved.[24] The correlation of the dynamic expression of a proteome and the physiological changes related to a healthy or diseased condition can help to:

- support the understanding of disease mechanisms,
- design new ways for the discovery and validation of
- disease models,
- find new diagnostic markers,
- identify potential therapeutic targets,
- optimise lead compounds for clinical development,
- characterise drug effects,
- Study protein toxicology.

Blackstock and Weir[15] proposed that pharmaceutical proteomics for target validation split into expression proteomics and cell mapping or interaction proteomics, each having a distinct role in the overall drug discovery process.

Cell-mapping proteomics has a more defined goal of studying protein-protein interactions by systematically characterising the components of protein complexes and building up a map of cellular pathways and interactions that may be important either in a disease process or in the mechanism of action of a drug. By use of specific antibodies or artificially introduced tags, specific proteins can be isolated and any associated proteins can be identified rapidly by mass spectrometry.

Expression proteomics is the profiling of protein expression in a cell under various stimuli, probably of most use in the search for surrogate markers of drug responsiveness and in toxicology. Such an approach has formed part of the Developmental Therapeutics Program at the US National Cancer Institute, where 3989 compounds have been screened against a panel of 60 cell lines, and their molecular pharmacology characterised by two-dimensional electrophoresis.[25]However, expression proteomics must compete with the genomic analogues of differential gene expression and chip-based hybridisation technologies. For example, with a DNA microchip, [26] the temporal expression changes in 8613 genes in human fibroblasts responding to serum was examined. Such technologies provide results more rapidly and are less labour-intensive than the corresponding proteomics-based approaches but, given the poor correlation between mRNA and protein product and the higher value of the protein information, substantial investment is being made to automate large-scale proteomics operations.[22]

Proteomics will play a major role in biomedical research and it will have a significant impact on the development of diagnostic and therapeutic products such as cancer, heart and infectious diseases in the future.[27]

Clearly, genomic and proteomic approaches complement each other in terms of the information produced and their relative advantages and disadvantages. Modern drug discovery is a highly competitive process. It is not sufficient to identify a disease-related protein target by proteomics. Required is the understanding of the biochemistry and the regulation of an appropriate protein pathway or cascade, in order to find the best possibility to interfere.

CONCLUSION

By contrast to the agents administered to patients in clinical wars, the process of drug discovery is not a prescriptive series of steps. The risks are high and there are long timelines to be endured before it is known whether a candidate drug will succeed or fail. At each step of the drug discovery process there is often scope for flexibility in interpretation, which over many steps is cumulative. The

pharmaceutical companies most likely to succeed in this environment are those that are able to make informed accurate decisions within an accelerated process.

The genomics revolution has impacted very positively upon these issues and now has a powerful new partner in proteomics. The ability to undertake global analysis of proteins from a very wide diversity of biological systems and to interrogate these in a high-throughput, systematic manner will add a significant new dimension to drug discovery. Each step of the process from target discovery to clinical trials is accessible to proteomics, often providing unique set of data. Using the combination of genomics and proteomics, scientists can now see every dimension of their biological focus, from genes, m RNA, proteins and their subcellular localization. This will greatly assist our understanding of the fundamental mechanistic basis of human disease and allow new improved and speedier drug discovery strategies to be implemented.

REFERENCES

1. J.Martin. Proteomics as a major new technology for the drug discovery process. Drug Discovery Today 1999; 4(2):55-62.
2. Sweeny K. Technology trends in drug discovery and development: Implications for the development of the pharmaceutical industry. Centre of strategic economic studies 2002; 3:1-28.
3. Bohacek R.S. The art and practice of structure-based drug design: a molecular modeling perspective. Med. Res. Rev. 1996; 16:3–50.
4. Pharma reports. Advances in genomics and proteomics transform drug discovery and development. Pharma and Biotech Industry Reports 2008; 46.
5. Grant S.G, Stock W.P. Difference between genomics and proteomics. J. Neuro Sci. 2001; 21(21): 8315-8.
6. Shine J, Suburg P.H, Martial J.A, Baxter J.D, Goodman H.M. Construction and analysis of recombinant DNA for human chronic somatotropin. Nature 1977; 270:494-9.
7. Bassett D.E, Boguski M.S, Spence F, Reeves R, Kim S, Weaver T, Hieter P. Genome cross-referencing and XRE fbd: Implications for the identification and analysis of genes mutated in human disease. Nature genet, 1997; 15:339-43.
8. Mekusick V.A. Genomics: Structural and functional studies of genomes. Drug Discov. Today 1997; 45(2): 244-9.
9. National Human Genome Research Institute. FAQ about genetic and genomic science. Genome. gov. 2010.
10. Susan J. Ward. Drug discovery and genomics technologies: Impact of genomics in drug discovery. Bio. Techniques 2001; 31(3).
11. Emilien G, Ponchon M, Caldas C, Isacson O, Maloteaux J.M. Impact of genomics on drug discovery and clinical medicine. QJM: An intrnational journal of medicine. 2000; 93(7):391-423.
12. Wang L. Pharmacogenomics: a systems approach. Wiley Interdiscip Rev Syst Biol Med. 2010; 2 (1): 3–22.
13. Issa, A.M. Pharmacogenetics, ethical issues: review of the nuffield council on bioethics report. Trends Pharmacol. Sci., 2000, 21: 247.
14. Anderson NL, Anderson NG. Proteome and proteomics: new technologies, new concepts, and new words. Electrophoresis 1998; 19 (11): 1853–61.
15. Blackstock WP, Weir MP. Proteomics: quantitative and physical mapping of cellular proteins. Trends Biotechnol. 1999; 17 (3): 121–7.
16. James. Protein identification in the post-genome era: the rapid rise of proteomics. Quarterly reviews of biophysics 1997; 30 (4): 279–331.
17. Marc R. Wilkins, Christian Pasquali, Ron D. Appel, Keli Ou, Olivier Golaz, Jean-Charles Sanchez, Jun X. Yan, Andrew. A. Gooley, Graham Hughes, Ian Humphery-Smith, Keith L. Williams & Denis F. Hochstrasser. From Proteins to Proteomes: Large Scale Protein Identification by Two-Dimensional Electrophoresis and Arnino Acid Analysis. Nature Biotechnology 1996; 14 (1): 61–65.
18. UNSW Staff Bio: Professor Marc Wilkins.
19. Yoshida M, Loo J.A, Lepleya R.A. Proteomics as a tool in the pharmaceutical drug design process. Curr. Pharm. Des. 2001; 7(4):291-310.
20. Pandey A and Mann M. Proteomics to study genes and genomics. Nature 2000; 405: 837-846.
21. Anderson N.L, Seilhamer J.A. A comparison of selected RNA and protein abundance in human liver. Electrophoresis 1997; 18:533-37.
22. Wilkins M.R, Sanchez J.C, Williams K.L, Hachstrasser D.F. Current challanges and futur applications for protein maps and post-translational, vector maps in proteome projects. Electrophoresis 1996; 17:830-38.
23. Druker B.J, Lydon N.B. Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukemia. J. ClinInvest 2000; 105:3-7.
24. Anderson N.L, Anderson N.G. New technologies, new concepts and new words. Electrophoresis 1998; 19: 1853-1861.
25. Myers T.G, Anderson N.L, Waltham M, et al. A protein expression database for the molecular pharmacology of cancer. Electrophoresis 1997; 18:647-53.
26. Iyer V.R, Eisen M.B, Ross D.T, et al. The transcriptional program in the response of human fibroblasts to serum. Science 1999; 283:83-87.
27. Pathade A.P, Bairagi VA., Ahire YS., Bhatia N. Proteomics: opportunities and challenges. Int J Pharm Sci Nanotech. 2011; 3(4):1165-1172.