THE ADVANCES OF PROTEOMICS

Os avanços da proteômica

Carolina Amaral Oliveira1
Daniel Vinicius Alves Silva1
Deborah Cristina Brito Sales1
Fernanda Aparecida Cardoso Mendes1
Patrícia Oliveira Silva1
Samara Frantheisca Almeida Barbosa1
Caio Roberto Soares Bragança2
Joanilha Ribeiro Lopes1

Abstract: To analyze the importance of proteomics in the identification and characterization of individual proteins, elucidating their associations and functions in health and disease, and the main proteomic techniques for these applications. It is a review of bibliographical literature, in which the inclusion criteria of the articles were: studies in the whole, in the Portuguese language, from 1999 to 2016 and that answered a guiding question of the research. In this search 122 titles were found where 06 articles were chosen that served as basis for the discussion besides other complementary literatures. Proteome is a relatively new term, and the study of this science led to the establishment of two main techniques with great application in the domains of life - two-dimensional electrophoresis and mass spectrometry. The characterization of proteins is a process, based essentially on the techniques of two-dimensional electrophoresis and mass spectrometry, which perform the separation and identification of proteins, respectively. We conclude that the contribution of proteomics to the systemic understanding of biological function in its molecular context has become an important tool in the scientific area. Thus, further research into proteomics is necessary, since studies of protein profiles can be extended to all three domains of life - Eukarya, Bacteria and Archaea proteomics. In the health context, in particular, this technology could provide people with differentiated, resolute and higher quality care.

Keywords: Proteome; Proteomics; Protein analysis; Two-dimensional electrophoresis; Mass spectrometry.

Corresponding author: Carolina Amaral Oliveira.
E-mail: carol_oliveira13@hotmail.com.

1 Universidade Estadual de Montes Claros.
2 Instituto Superior de Educação Ibituruna.
Resumo: Analisar a importância da proteômica na identificação e caracterização de proteínas individuais, elucidando suas associações e funções no âmbito da saúde e doença e as principais técnicas proteômicas para essas aplicações. Trata-se de uma revisão de literatura, em que os critérios de inclusão dos artigos foram: estudos na íntegra, no idioma português, no período 1999 a 2016 e que respondessem questão norteadora da pesquisa. Nesta busca foram encontrados 122 títulos dos quais foram escolhidos 06 artigos, que serviram de base para a discussão, além de outras literaturas complementares. Proteoma é um termo relativamente novo, e o estudo desta ciência levou o estabelecimento de duas técnicas principais com grande aplicação nos domínios da vida - a eletroforese bidimensional e a espectrometria de massa. A caracterização das proteínas é um processo, baseado, essencialmente, nas técnicas de eletroforese bidimensional e espectrometria de massas, que realizam a separação e a identificação das proteínas, respectivamente. Conclui-se que a contribuição da proteômica para a compreensão sistêmica da função biológica em seu contexto molecular tornou-se ferramenta importante na área científica. Assim, novas pesquisas voltadas para a proteômica são necessárias, uma vez que os estudos dos perfis proteicos podem ser estendido a todos os três domínios da vida – Eukarya, Bacteria e Archaea proteômica. No contexto da saúde, em particular, essa tecnologia poderá proporcionar às pessoas uma assistência diferenciada, resolutiva e com maior qualidade.

Palavras-chave: Proteoma; Proteômica; Análise de proteínas; Eletroforese bidimensional; Espectrometria de massa.
INTRODUCTION

In the last decades there has been a rapid development in studies of the sequencing of genomes of various organisms. Therefore, a large proportion of molecular information about the biological systems is available for public consultation.¹

This rapid and increasing availability of genomic data has awakened great scientific interest, however, the accumulated knowledge through the results of DNA sequencing from different organisms has shown that it is still necessary further studies oriented in an attempt to understand them in a certain physiological condition. However, genomic sequences continue providing expertise in areas as diverse as medicine and human evolution.¹

In fact, the genome projects involve mapping, sequencing, the analysis and comparison of genomes of a whole organism or any specific region, not revealing, thus data on the expression of their genes, the quantity and the functionality of their products. The study of the genome does not allow to identify which proteins are expressed in a particular cell in a given situation. Within this context, it is important to study on a large scale of proteins through projects of proteomic analysis, which will allow, in this way, to investigate whether a gene is being expressed, the relative concentration of this product, and finally, the modifications that may occur in these proteins after their translation.¹²

The term proteome was introduced in 1995 to describe the number and types of proteins present in an organism. It is known that proteins are responsible for controlling most of cellular processes, which occur in great diversity, acting as enzymes, antibodies, hormones, structural components and cellular receptors.³⁵ Due to the dynamic nature of the proteome subject to change in response to the environment of the organism or other factors, such as the development cycles, their study represents a way to investigate possible functions of proteins in a given time.¹

The data generated by proteomics analysis allow to achieve different objectives, as for example, (1) clarify the proteins involved in metabolic pathways related to various cellular processes; (2) identify new pharmacological targets and biological markers related to the process of establishment and progression of diseases; (3) identify bioactive molecules from biological natural extracts, leading to the development of new drugs; in addition to (4) characterize the cellular responses to certain drugs, diseases and environmental changes.¹

Therefore, researchers from all over the world have worked in several lines of research involving expression of proteins, such as the effect of drugs, pathological conditions, cellular differentiation, comparison of varieties of the same species and cellular responses to external stimuli, and they may, in this way, to easily identify the proteins expressed or repressed, through the achievement of two-dimensional profile of proteins from their system.⁵ Thus, instead of analyzing separately each of the protein components of a system, the proteomics aims to contribute to the systemic understanding of biological function of proteins in its complex molecular context.

Currently, there are several techniques for the study of proteomics, as two-dimensional electrophoresis in polyacrylamide gel electrophoresis (2D) and mass spectrometry. The first important approach of proteomics arose with beginning of 2D. Its main advantage is the ability to separate, with high resolution, a large number of proteins present in a cellular sample.
cell, in addition to the possibility of carrying out the analysis of gene expression by comparing the protein patterns.1,6 More recently, proteins were identified by mass spectrometry, usually after the preliminary digestion, to generate a characteristic set of peptides. Thus, all the information generated may be sufficient to identify completely the protein.7

Before this awaking of post-genomics era, there appeared the following question: what is the importance of new technologies generated by proteomics in the health area?

Therefore, the objective of this study was to analyze the importance of proteomics in the identification and characterization of proteins, explaining their associations and functions in the context of health and disease.

It is about a literature review. The bibliographic survey was carried out in the access network of the Virtual Library of Health (VHL), in electronic journals (Biotechnology Science and Development; Reviewed Encyclopedia Biosphere; Health & Environment Review; Journal of Intensive Care Unit). Articles were selected using as descriptors: proteome, proteomics, post-genomics, two-dimensional electrophoresis, mass spectrometry.

The inclusion criteria of the articles were: full articles, in Portuguese Language, in the period from 1999 to 2016 and which answered to the main question of the research. The following articles were excluded: those found in more than one source of information or duplicated, those not directly related to the theme and, finally, the dealing of the subject, but did not comply with the needs of the authors.

In this search 11 titles of BVS were found together with the journals aforementioned. For the initial selection of articles, reading was performed of all titles and summaries found, considering the pre-established inclusion and exclusion criteria, from this step 06 articles were pre-selected as a basis for discussion and other complementary literatures.

Thereinafter, it was followed to step of treatment of results and interpretation, highlighting the relevant information that emerged from the categories, making interference and interpretations based on the theoretical material which this process was based on.

LITERATURE REVIEW

Table 1 shows the selected articles on the topic of revision according to title, authors, results, conclusions and recommendations in the period from 1999 to 2016.

As for the objective of publications, the authors sought, in their majority, to talk about the importance of specific study of proteins, i.e., investigate or clarify their possible roles in the metabolic processes of living systems for a better understanding of the functioning of a cell or tissue at the molecular level.

Table 1 - Distribution of articles located at BVS

<table>
<thead>
<tr>
<th>Article Title</th>
<th>Authors</th>
<th>Results</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Proteomics - a functional approach to the study of the genome.</td>
<td>SILVA, A.M.S et al.1</td>
<td>The genomic sequencing and the consequent knowledge of the full sequence of all genes are important contributions to the study of better living organisms. However, such information is not enough to know which proteins are being expressed by the cell, at a given moment, in a given condition. Due to the dynamic nature of the genome, their study represents a way to search possible functions of proteins and a way to investigate metabolic processes in living systems to better understand the functioning of a cell or tissue at the molecular level.</td>
<td>The growing generation of genomic data and technological progress in the study of proteins made the proteomic analysis, became a powerful tool for studies on physiology and genetics of various living organisms, including plants. In the medical area several proteomic studies were performed in the characterization of cancer, Alzheimer’s disease, infectious and cardiac diseases, differentiation of infectious agents, non-pathogenic pathogenic strains that can assist in the development of diagnostic methods and therapeutic agents.</td>
</tr>
</tbody>
</table>
#### Article Title

<table>
<thead>
<tr>
<th>Authors</th>
<th>Results</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2. Proteomes Analysis</strong>&lt;br&gt;SILVA, A.M.S et al.¹</td>
<td>In the proteomes projects, one of the primary goals is to separate and visualize the maximum possible proteins from one source, allowing them to be cataloged computationally and studied by analytical techniques. Currently the two-dimensional electrophoresis in polyacrylamide gel electrophoresis is the most efficient method of simultaneous separation of hundreds or thousands of proteins. It can be said that the two-dimensional electrophoresis is the core of proteomes analysis.</td>
<td>It can be concluded that it is of great importance the existence of available maps of Proteomes, because they allow researchers from all over the world, working in various matters involving expression of proteins, such as the effect of drugs, pathological conditions, cellular differentiation, comparison of varieties of the same species and cellular responses to external stimuli, can more easily identify the expressed or repressed proteins, needing to do this just to get the profile of bi-dimensional proteins from its system under the same conditions of maps of Reference available.</td>
</tr>
<tr>
<td><strong>3. Proteomes: recent advances in techniques of bi-dimensional</strong>&lt;br&gt;CIERO, L. D; BELLATO, C. D. M.8</td>
<td>Proteomics can be seen as a methodology for selection of molecular biology, which has as its objective to document the general distribution of proteins of the cell, identify and characterize individual proteins of interest and especially to elucidate their associations and functions. There are several techniques, but the one that has greater relevance is the 2DE, is a method of efficient separation, because all proteins in a sample are separated simultaneously, providing useful information about IP, molecular, expression and relative abundance and post-translational modification by change of the electrophoretic mobility.</td>
<td>The proteomes came to revolutionize the study of Genome, once it uses tools to identify the proteins expressed in specific situations at a given moment in that environment. Proteomics is dynamic and allows to perform comparative studies. The technological innovations in the detection, analysis and quantification of proteins has accelerated rapidly the development of proteomics in the past 4 years.</td>
</tr>
<tr>
<td><strong>4. Proteomics: New frontiers at the clinical research.</strong>&lt;br&gt;RIVEROS. A.C.G. et al.⁹</td>
<td>The perspective offered by proteomics has been used in research on different areas of medicine, including e biomedicine. These surveys could be classified in different ways: depending on the type of sample used, the disease or the type of diseases that they address, technique or techniques used, the use or application. Studies have shown that the number of proteins identified in each proteome varies significantly, which is probably a reflection of the existence of differences in the level of expression of the different proteins in each cellular line. Human tumor cell line may constitute an important model for the study of cancer.</td>
<td>The proteomics aims to compare, at the molecular level, the cellular protein level cellular in normal and pathological situations. The study of mass spectrometry of proteins separated by bi-dimensional gel is leading to a rebirth in biochemical approaches about the function protein. Major advances in clinical research have been developing using these proteomic tools, which will be the key to new discoveries of drugs and treatments of diseases.</td>
</tr>
<tr>
<td><strong>5. Bi-dimensional electrophoresis and mass spectrometry as proteomic tools applies to the definition of protein markers associated to the reproductive efficiency of goats.</strong>&lt;br&gt;JUNIOR, R. Q. B, et al.¹⁸</td>
<td>The fact of the seminal plasma play an important role in the maturation of sperm makes it a biological material of common use in many works of proteomics, aiming to determine protein profile, i.e., characterize and identify proteins such as, for example: the proteins of the sperm (Binderspermproteins-BSPs), matrix metalloproteinasises (MMPs); proteins with affinity to heparin (HAPs); spermadhesins; phospholipase A2 and lactoferrin.</td>
<td>The use of analytical techniques for proteomics, as the two-dimensional electrophoresis (2DE) and mass spectrometry, applied to the identification of proteins related to the reproductive capacity of the goats herd, contributes to the growth and development of the reproduction and improvement of zoo technical potential within the herd. The achievement of results like these helps in the growth and development of the reproductive stage because it will enable a greater understanding and identification of differences of potential fertility of animals.</td>
</tr>
</tbody>
</table>

---

Continuação da table 1.
The proteomics came to transform the Genome studies, once that its tools are used to identify the proteins expressed in specific situations at a given moment in that environment.\(^5\) It is seen as a methodology for selection of molecular biology, which has as its objective to document the general distribution of the cell proteins, identify and characterize individual proteins of interest and especially their functions. The technological innovations in the detection, analysis and quantification of proteins has accelerated rapidly the development of proteomics in the last years.\(^8\)

The proteomic projects have as main objective the study on a large scale, or in genomic scale, structure, function and regulation of proteins of an organism, thus allowing that they are catalogued computationally and studied by analytical techniques. The perspective offered by proteomics has been used in research on different areas of health, being based on the type of sample used as the research can be highlighted with cell lineages, with tissues and proteomics of fluids.\(^5,9\)

The study of proteome comprises the establishment of some techniques, the most commonly used is the two-dimensional electrophoresis (2-DE) in conjunction with the mass spectrometry (MS). The first is based on the separation and detection of proteins, while EM is applied in its identification.\(^10\)

The 2-DE was introduced in the 1970s and the steps that make it consist of: Sample preparation; separation by isoelectric point; separation by size; detection; analysis of gels images; selection of proteins of interest; cut of sections of the gel with selected proteins; digestion of proteins with a specific protease; analysis of mixtures of peptides by mass spectrometry; and identification by comparison of proteins databases.\(^11\)

On the other hand, mass spectrometry is a mechanism for accurate determination of molar weights, which allows to obtain information that allows for the resolution of various issues related to the chemical structure of proteins, such as: checking the correctness of a sequence of amino acids, determination of the fidelity and consistency of recombinant proteins, detection of genetic diseases, among many others.\(^12\)

This is all made possible because the
The technique of proteomics provides, in addition to the molecular, structural characterization of proteins, the sequencing of peptides and post-translational modifications. EM stands out especially in qualitative analyzes, such as the identification of compounds in mixtures and structural characterization of unknown compounds.\(^{13-14}\)

In a study done in animals with the techniques of two-dimensional electrophoresis and mass spectrometry, proteins were found similar to those of connection to sperm, upon evaluating the protein profile of seminal plasma of sheep.\(^{15}\) The results found when examining the seminal plasma of dogs, suggested that the increased activation of pro-MMP-2 can benefit the viability and motility of sperms.\(^{16}\) In one study, it was observed that the pattern of HAPs, proteins with affinity to heparin, has relation with the season of the year. In fact, all these findings allowed a better understanding about the role played by these proteins.\(^{17}\) In addition, they showed once more the importance of these techniques in different areas of study.\(^{18}\)

Although the 2D is able to generate a lot of information, some limitations can also be observed when very acidic proteins (pH lower than 3.5) or basic (pH higher than 9), proteins of low abundance and hydrophobic protein, which are usually found in bio membranes, are subjected to analysis.\(^{8}\)

In addition, the proteome can be considered dynamic due to the physiological status and the stages of cellular differentiation that the organism is. It is estimated that more than a million different kinds of proteins are present in the cells, tissues and body fluids under distinct conditions and/or times.\(^{19}\)

Despite of these challenges, there are no doubts that the results of the proteomic approaches are potentially useful in many areas of clinical research, among them, diagnosis, monitoring of response to therapy, prediction of clinical outcome, classification of subtypes of diseases, determination of risks, characterization of the metabolic pathways, quantification of biomarkers and generation of therapeutic targets.\(^{19}\)

The growing generation of genomic data and the technological advancement in specific knowledge about the structure and function of proteins, meant that the proteomic analysis has become a powerful tool for studies on the physiology and genetics of various living organisms.\(^{1}\) Due to being a new study, these techniques and methodologies are still being refined to obtain relevant information about the normal and abnormal processes that occur in the human body.\(^{8,19}\)

The perspective offered by proteomics has been used in research on different areas of medicine, including biomedicine. In the medical field, several proteomic studies were performed in the characterization of cancer, neurological diseases such as Alzheimer’s disease, infectious diseases and heart diseases. Other studies show the characterization of infectious agents such as *Mycobacterium tuberculosis* (TB), including comparing the proteome of pathogenic and non-pathogenic strains of microorganisms, which may assist in the development of diagnostic methods and therapeutic agents.\(^{1,8}\)

These surveys about proteomics, could be classified in different ways: depending on the type of sample used, the disease or the type of diseases that they address, technique or techniques used, the use or application, etc.\(^{8-9}\)

When it comes to studies comprising cellular lineages, the proteomics has determined the profiles of proteins expression in different types of these lineages which in the future will be used as the basis for comparative experiments. Studies have shown that the number of proteins identified in each proteome varies significantly, which is hypothetically a reflection of the existence of differences in the level of expression of the different...
proteins in each cellular lineage.

The analysis of tumor cell lineages, for example, constitute an important form of study on cancer. Therefore, researches have been developed by comparing the proteomes of human tumor lineages, which can assess the malignant potential of a tumor, besides its chemiosensitivity. This analysis has become very important for an early detection, classification and choice of prognosis of tumors, contributing to new modalities of treatment.

In addition to the research in cell lineages, the proteomics has currently invested time in the analysis of tissues and fluids for the study of physiological situations (during development, in different metabolic stages, forward the various environmental responses) or pathological (cancer, autoimmunity, infection), being similar to studies with cell lines.

However, it is debatable the difficulty of the examination of affected tissues, since in the majority of diseases in humans, they are not easy to access and are unlikely to be used for analysis. Thus, there is cellular heterogeneity, which can lead to inaccurate results if a detailed histopathologic evaluation is not performed. In order to overcome this difficulty, there is a laser microdissection, however it is characterized by the generation of a small number of cells and introduction of an extra step in the processing of the sample.

In fact, the inaccessibility of tissue samples is a problem, because the information obtained through these, can be decisive and, sometimes, it is the best choice for a possible treatment of a disease.

The proteomics of fluids presents a greater ease of obtaining and processing of samples. A type of sample used by means of this technique is through the proteomic analysis of cerebrospinal fluid. This has been used recently in the study of etiopathogenic basis and in the identification of biomarkers for neurological diseases, such as neuropsychiatric disorders, brain tumors and painful syndromes of the lumbar spine.

In addition to the cerebrospinal fluid, the parameters of saliva flow and composition have been exploited in the monitoring of hormone levels and drugs, exposure to environmental pollutants and infectious, and monitoring of diseases, including periodontal disease, diabetes mellitus, cystic fibrosis, Sjogren’s syndrome, Diseases of the salivary gland and breast cancer, ovarian cancer and oral cancer.

In Brazil, despite of the own concept of genomics be little unexplored by the scientific community and funding agencies, proteomic projects began to be performed in the Laboratory of biochemistry and chemistry of Proteins/Brazilian Center of Services and Research in Proteins (CBSP/LBQP) at the University of Brasilia (UnB). CBSP/LBQP has been dedicated in recent years to identify proteins through analysis of amino acids, automatic sequencing and mass spectrometry, having acquired the necessary know-how for the work in cellular proteomes.

One of the projects of LBQP/CBSP is through the analysis of genomics of human leukocytes after trauma, in collaboration with Doctor Belchor Fontes, of Hospital das Clínicas-São Paulo city. After the trauma, even in not very serious cases,
some patients develop multiple organ failure, and they may even die. Thus, leukocytes of patients with and without multiple organ failure after trauma can be compared at the level of proteomes, with the aim of identifying molecular markers associated with this condition.

The proteins, in a general way, control the life and health of a body, therefore, the objective of this research was to understand how proteins operate in a biological system in order to understand how to regulate the mechanisms of disease to a further treatment. It was still not possible to achieve maximum effect of proteomic approaches in biomedical research, in part, because of the lack of resources and information among the researchers on the technological advancements.

The analysis of proteomes has many other applications beyond those cited, in this literature review, and the demand for work in this field will certainly grow exponentially in the coming years. The care of this demand will depend on the continuous support to research groups in the study of human proteome, thus increasing the scale of measurement of large groups of proteins, allowing for a better and integrated overview of pathological and physiological processes, by means of analyzes of protein profiles.

FINAL CONSIDERATIONS

The proteomics is a modern field of studies that has contributed, significantly to the understanding of systemic biological functions in its molecular context. Due to being a dynamic process, with vast complexity and diversity of information, it becomes a necessary and revolutionary tool in the scientific area. In the context of health, it has already led to the generation of new methods of diagnosis, monitoring of response to therapy, classification of subtypes of diseases, determination of risks, quantification of biomarkers and generation of therapeutic targets.

The improvement of techniques for two-dimensional electrophoresis and mass spectrometry are paramount to the continued advancement in the study of proteome, still lacking in studies. New bodies of research, with a focus not only on the clinical area, but throughout the field of proteomics need to be stimulated, allowing the discovery of information and the growth and development of various sectors, since the studies of protein profiles can be extended to all three domains of life - Eukarya, Bacteria and Archaeproteomics. In the health context, specifically, this technology shall provide the people with a specialized, resolute, support and with better quality.

This is a non-sponsored study. There are no conflict of interests.

REFERENCES


