Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/dit

Review Article

Proteomics & metabolomics: Mapping biochemical regulations

Anjum Gahlaut [a,*], Vikas [a], Monika Dahiya [a], Ashish Gothwal [a], Mahesh Kulharia [b], Anil K. Chhillar [a], Vikas Hooda [a], Rajesh Dabur [c]

[a] Centre for Biotechnology, Maharshi Dayanand University, Rohtak, Haryana 124001, India
[b] Centre for Bioinformatics, Maharishi Dayanand University, Rohtak, Haryana 124001, India
[c] Department of Biochemistry, Maharishi Dayanand University, Rohtak, Haryana 124001, India

ARTICLE INFO

Article history:
Received 28 June 2013
Accepted 27 August 2013
Available online 20 November 2013

Keywords:
Proteomics
Metabolomics
Chromatography
Spectrometry
Bioinformatics

ABSTRACT

Large-scale studies in the field of omics have been successfully exploring the differences in gene expression, protein and metabolite abundance and modification of post-translational protein, and providing a different level of views for the cellular processes occur in cells. Proteomics and metabolomics are new addition to the 'omics' field, but both of them are still developing its own computational infrastructure by assessing the computational needs of its own. Due to the strong knowledge on chemical information and the importance of linking this chemical information to biological consequences, proteomics and metabolomics combines the elements of traditional bioinformatics and cheminformatics.

1. Introduction

The study of biological entities at the system level is a clear trend in the life sciences. Analytical tools are required to identify the component parts of the system and determine their responses to a changing environment. To achieve all these requirements, a combination of transcriptomic, proteomic, and metabolomic profiling technologies have been developed, and among these technologies, proteomics is continuing to evolve rapidly. Presently, there are large numbers of proteomic studies have been published in the literature, only a small portion has attempted to provide an extensive quantitative description of the biological system under investigation. Apart from the phenomenal contribution of the mass spectrometry and peptide separation techniques in area of proteomics studies, there is so many unsolved technical challenges for identification and quantification of all of the proteins in the biological system is still remain. While proteomic data for the genome of unicellular organisms has been occasionally achieved beyond 50% but the proteomic coverage for multicellular or higher organisms strictly exceeds more than 10%. For protein quantification, these figures have low data quality, in terms of available information content, because the information required for quantification are more than for protein identification.

* Corresponding author.
E-mail address: anjumgahlaut@gmail.com (A. Gahlaut).
0975-7619/$ – see front matter
http://dx.doi.org/10.1016/j.dit.2013.08.007

The use of fluorophores, dyes or radioactivity in classical proteomic quantification methods provides very good linearity, sensitivity and dynamic range, but they have two important drawbacks: (1) requirement of protein separation at high resolution which is typically provided by 2D gels, so can not be applicable to abundant and soluble proteins, and (2) they do not provide the information related to the underlying protein. Both of these problems can be solved by using the modern LC-MS/MS techniques. However, mass spectrometry is not used for quantitative purpose due to the wide range of physicochemical properties such as size, charge, hydrophobicity etc. exhibited by proteolytic peptides; this causes the large differences in mass spectrometric response. The accuracy in protein quantification can be achieved by comparing each individual peptide between experiments.

In order to achieve a complete analysis of the biological response of a complex system, it is important to monitor the response of an organism to a conditional difficulty at the transcriptome, proteome and metabolome levels.[1] Integration of experimental data with the results of functional genomics is an important step to achieve this goal. Metabolomics can be considered as the most recent contribution to this area. It involves the qualitative and quantitative analysis of all the metabolites in the cell (the metabolome). Moreover it is more closely related to the organism's actual phenotype and can be linked to the genotype through the knowledge provided by the biochemical pathways and gene regulatory networks.[2] Comprehensive studies of metabolic processes have been made possible and useful with the development of modern analytical and computational tools.

While transcriptomics and proteomics studies provide critical insight into sequential modulation of metabolic reaction flux but metabolomics may provide information related to regulation called, metabolic regulation.[3] The metabolic regulation can be described as the effect of metabolite concentrations on actual activity of enzyme through mass action, kinetic and allosteric effects.[4]

2. Proteomics

Proteomics is the simultaneous and systematic analysis of the diverse nature of proteins. The aim to develop proteomics is to provide detailed information about the structure and function of biological systems in different biological conditions. A proteome may be defined as the total content of proteins expressed by a genome in a cell or tissue at a certain time. The term proteome was first introduced in the 1990s.[5,6] Analysis of proteome is most commonly performed by a combination of two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS). With the help of 2-DE technique, a complex and variable mixture of protein is separated and visualized and then for identification of protein of interest, mass spectrometry is applied. The proteome of one organism differs from other organisms, depending on the genome and on external and internal factors including health, stress, physiological state, disease and drugs. The complexity of the proteome is far higher as compared with the genome because of the protein processing and modification. The main focus of proteomics studies is to provide detailed descriptions of the diverse properties of proteins in a variety of biological systems.[7] Although proteomics is a relatively new field, but the new methodologies in the proteomics studies have been under development for decades.

Proteomic study of proteins is generally based on four technological parameters, (i) a simple and fast method for purification of proteins in small amounts from complex mixtures, (ii) a rapid and sensitive method to generate sufficient detailed structural information for protein molecule being studied, (iii) access to structural and sequence databases of protein or DNA, and (iv) computer-based algorithms capable of translating and linking the language of DNA sequence with various types of structural information of protein like internal peptide sequences or N-terminal protein, composition of amino acids, p/peptide mass fingerprints, sequence tags of selected peptides or mass spectrometry fragmentation patterns.

2.1. *Proteomics methodologies*

The proteins separation on the whole protein level is usually performed by gel-based electrophoretic or by liquid chromatographic methods. Peptide level separations or fractionations can be achieved by chromatographic methods or by peptide isoelectric focusing.[8–10]

2.1.1. *Separation by gel-based method in proteomics*

After obtaining the fraction of desired protein by purification, this fraction is subjected to one-dimensional gel electrophoresis (1-DE) for resolving the relatively simple protein mixtures. In 1-DE, the separation of proteins according to their molecular weight is the basis of this method. 2-DE is used as a standard gel-based separation method in proteomics, which enabling the simultaneous separation and visualization of thousands of proteins in one time. In 2-DE, isoelectric focusing (IEF) is used to separate proteins in first dimension according to their isoelectric point (pI) in a pH gradient, and after that proteins are separated according to their molecular weight in second dimension.

2.1.2. *Separation by non-gel-based method in proteomics*

The introduction of non-gel-based strategies in the field of proteomics has provided high-throughput methodologies. The emergence of non-gel-based proteomic methods in recent years is due to the availability of proteins separation by liquid chromatographic techniques, new protein chemistry, new enrichment methods and the development of mass spectrometry and new software for data analysis. Quantitation by mass spectrometry provides important addition to quantitation by 2-DE. The use of mass spectrometric-based technologies have several advantages over 2-DE-based technology as they are automated and separated complex peptide mixtures with high resolution and high sensitivity.[11] To perform liquid chromatography/mass spectrometry (MS) in proteome analysis, this require a complexity reduction in order to detect and analyze many possible components in the sample.[6] This can be achieved by combining the two different orthogonal peptide separation methods such as cation-exchange chromatography and capillary reversed phase chromatography in combination with MS/MS. The