

Proteomics: Yesterday, Today and Tomorrow

Ida Pucci Minafra *

Centro di Oncobiologia Sperimentale, Palermo, Italy

***Corresponding author:** Prof. Ida Pucci Minafra, Centro di Oncobiologia Sperimentale, Via San Lorenzo Colli 312, 90146 Palermo, Italy, Email: pucci.ida@gmail.com

Commentary

Volume 1 Issue 3

Received Date: November 12, 2017

Published Date: December 07, 2017

Commentary

The overwhelming entry of proteomics into the scientific world [1], catapulted scientists into the era of bio-complexity of the living matter. Researchers have always been self-questioning about the essence of life, often receiving inadequate and frustrating responses. Unexpectedly, since the middle of the last century, the perspective of research has changed profoundly: a "Copernican revolution" was raised by the publication of the DNA structure in 1953 by Watson and Crick, as is well known. The long wave of this discovery influenced research for many years. It was thought to have discovered the secret of life and that the genes were the absolute protagonists of creation, reproduction, and biodiversity of the species. Subsequently, in 1970, Crick coined the central dogma of biology "a gene-a RNA-a protein", which confirmed the hypothesis of "an enzyme-a gene" formulated by Horowitz and Leupold in 1951.

At that time, the proteins were essentially assigned a structural function, and, in some way an auxiliary role in the living scenery, because the leadership belonged to the genes.

The feeling that the researches on the genes were of the first class compared to the proteins or other fields was rooted in a subtle and non explicit way, in the scientific-academic world.

Another less striking, but equally significant revolution, was the discovery that the genes need protein control to function, as firstly formulated by Jacob and Monod, Nobel Prize in Medicine in 1965, and fully demonstrated until the modern epigenetics.

As previously mentioned, a historical turning point for the reevaluation of the role of proteins in the living matter was determined by the birth of proteomics, which owes its official entry into the scientific world to the application of two-dimensional electrophoresis on large sheets of gel (2D-IPG).

In the world of research, two conflicting feelings were derived from these first applications: on the one hand, the awareness of the power of the system that for the first time allowed researchers to highlight a variety of proteins in the same assay, determining two fundamental properties of each, the MW and the pI; on the other hand there was a disappointment for a poor quality of the gel that blurred the results. This limitation was overcome with the improvement of instruments and techniques, among which the introduction of first-dimension strips with a precasted pH gradient [2], and the introduction of mass spectrometry for in-gel digested proteins from 2D-IPG maps.

These applications have been able to largely standardize the techniques, allowing the researchers to create interactive networks and to establish numerous databases. Concurrently, the decoding of proteomic maps by dedicated software generated enormous initial enthusiasm, especially in the hope of identifying clusters and "constellations" of proteins capable of outlining different, normal and pathological, cell phenotypes. The era of individual markers was fading away with the knowledge that genes and proteins do not work alone, but form complex functional networks, as the databases dedicated to protein-protein interactions soon demonstrated. Within a few years, the scientific field of proteomics has been greatly expanded and diversified, i.e: qualitative proteomics, structural proteomics, functional proteomics, quantitative proteomics/profiling,

comparative proteomics, metabolomics, immunopeptidomics, cell organelles-subproteomics, etc, all supported by robust informatic systems. Certainly, these researches have greatly contributed to our knowledge of a cell phenotype and of its possible functions and dysfunctions. Interesting correlations have been found between over-expression or down-expression of certain proteins in cases of important pathologies such as cancer or hematological diseases.

At the same time, some problems in the system have emerged: some of a technical nature (e.g. limitations due to the low solubility or to the low amount of certain proteins) and others of a biological sense: the first of which is that cellular proteins of normal and diseased cells are, to a large extent, qualitatively overlapping. Rather, the protein expression levels may quantitatively vary, up to the appearance or disappearance of certain protein members in pathological cells compared to the normal counterpart. Another critical point of biological significance is that the dogma of one gene-one protein has collapsed, and perhaps also the hypothesis of "moonlighting protein" or "gene sharing" [3] that postulated the existence of numerous functions of the same protein as the primary gene product. In the light of the new discoveries of proteomics, these assumptions should be revised especially in relation to the identification an ever-increasing number of isoelectric variants of the same protein.

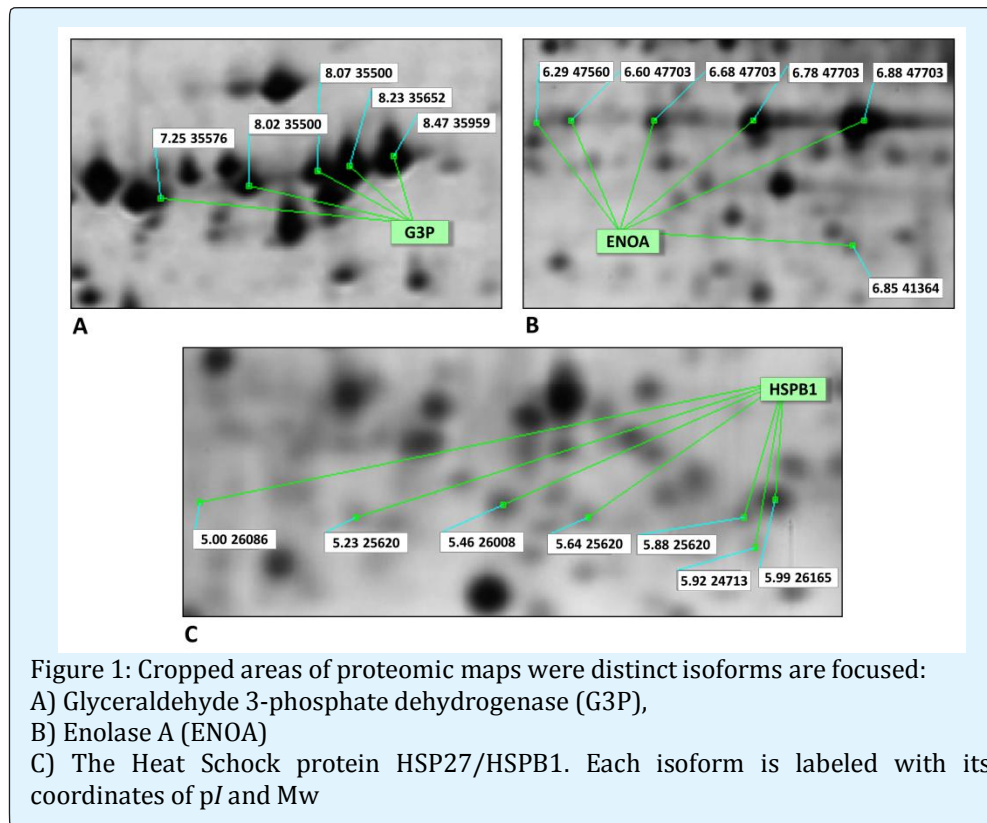
At present, in our laboratory, we have identified a number of isoforms and short forms in both the normal and the tumor breast proteomics, consisting of about 40% of the protein spots detected on a proteomic map [4]. The nature of these variants falls into two major categories: 1) isoelectric variants to be attributed to post-translational modification (i.e., glycosylation, phosphorylation, methylation, nitrosilation etc.) of the primary gene product, involving more or less significant changes of the pI without any perceptible variation of the MW; 2)

variants depending on post-transcriptional modifications that may result in different polypeptide length and hence in the MW. In many cases, a series of multiple close spots "trains of spots" is observed in the 2D-IPG maps.

One of the best known examples is that of the Glyceraldehyde 3-phosphate dehydrogenase (G3P) that, in our maps, appears as a set of five isoforms with distinct pI and equal MW (Figure 1A). Another well known example is the Enolase A (ENOA) which shows four isoforms and at least two short forms (Figure 1B).

Molecular researches have shown that many glycolytic enzymes, over-expressed in cancer for the well-known "Warburg effect", in addition to their canonical metabolic functions, can play many other cellular roles that support proliferation and/or tumor progression [5]. These additional functions of the enzyme are conditioned by the cellular compartments where the enzymes operate. Indeed, the G3P can translocate from its original cytoplasmic location into the nucleus or into the mitochondria where it may perform additional functions, among the most remarkable is the interaction with cyclin B-cdk1 and the promotion of cell proliferation [6]. Similarly, ENOA has been defined as a typical moonlighting protein [7,8] performing multiple functions at distinct cellular sites through the "gene sharing" mechanism [9].

Another interesting example of proteins with multiple isoforms are the heat shock (HSPs) family members. It has been reported that HSPs display elevated expression levels in many forms of cancer, where they can perform anti-apoptotic activities both spontaneous and generated by therapy [10]. Figure 1C shows a typical 2D-IPG separation of the HSP27/HSPB1 in a proteomic map of breast cancer tissue, which segregates into seven distinct isoforms with similar MW (G. Di Cara, R. Musso and I. Pucci Minafra work in progress).



In conclusion, most of the proteins once considered immutable instruments of the cell and tissue scaffold have turned out to be flexible and extraordinarily multifunctional elements capable of performing disparate functions in the cell, migrating into distinct compartments and interacting with numerous and different partner by generating network arrays, with distinct nodes depending on the functional target.

Future Perspective

Currently, two major objectives are pursued by the proteomic approach, which has demonstrated to be a very powerful instrument to study phenotypic profiles and biodiversity, in normal and diseased cells and tissues. The first is to gain more insights into the bio-complexity, through basic research and knowledge-sharing between the various scientific fields, genetic, molecular biology, cytology etc. The second regards the proteomic transfer into the clinical applications, i.e. Mardamshina M, et al. [11].

In the last decades several new approaches derived from the traditional proteomics, have been done. Among the

most popular, is the “shotgun proteomics”, a non-gel based proteomic strategy, which offers many advantages in speed and automation, over the gel-based techniques. Proteins are extracted from a biological sample and directly digested with a protease: the deriving peptides are separated by charge and hydrophobicity for the MS analysis. An obvious limitation of the non-gel based techniques is the loss of information on two fundamental properties of the proteins: *pI* and *MW*, which however are not essential for clinical applications.

Additional complementary proteomic approaches, are based on the differential labelling of protein extracts with stable isotopes, among which two approaches are currently most used: the SILAC for cells in culture (*Stable Isotope Labelling by Aminoacids in Cell culture*) and the ICAT (*Isotope-Coded Affinity Tag*) based on the incorporation of isotopic tags after protein extraction. Others, such as iTRAQ (*Isobaric Tag for Relative and Absolute Quantitation*), TMT (*Tandem Mass Tag*), or SISCAPA (*Stable Isotope Standards and Capture by Anti-Peptide Antibodies*) are used for relative quantification of a variety of sample types.

All these "system-wide" strategies represent a powerful instrument to re-write the molecular basis of many cellular processes and, hopefully, to understand their pathological deviations. For all these considerations, proteomics is projected in the future as a true scientific branch, rather than a mere methodological application.

References

1. Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, et al. (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (NY)* 14(1): 61-65.
2. Görg A, Obermaier C, Boguth G, Weiss W (1999) Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: Wide pH gradients up to pH 12, longer separation distances and simplified procedures. *Electrophoresis* 20(4): 712-717.
3. Jeffery CJ (2015) Why study moonlighting proteins? *Front Genet* 6: 211.
4. Pucci-Minafra I, Di Cara G, Musso R, Cancemi P, Albanese NN, et al. (2017) Retrospective Proteomic Screening of 100 Breast Cancer Tissues. *Proteomes* 5: 3.
5. Lincet H, Icard P (2015) How do glycolytic enzymes favour cancer cell proliferation by nonmetabolic functions? *Oncogene* 34(29): 3751-3759.
6. Carujo S, Estanyol JM, Ejarque A, Agell N, Bachs O, et al. (2006) Glyceraldehyde 3-phosphate dehydrogenase is a SET-binding protein and regulates cyclin B-cdk1 activity. *Oncogene* 25: 4033-4042.
7. Díaz-Ramos A, Roig-Borrellas A, García-Melero A, López-Alemay R (2012) α -Enolase, a multifunctional protein: its role on pathophysiological situations. *J Biomed Biotechnol* 2012: 1567958.
8. Jeffery CJ (1999) Moonlighting Proteins. *Trends Biochem Sci* 24(1): 8-11.
9. Jeffery CJ (2003) Multifunctional proteins: Examples of gene sharing. *Annals of Medicine* 35(1): 28-35.
10. Wu J, Liu T, Rios Z, Mei Q, Li X, et al. (2017) Heat shock proteins and cancer. *Trends Pharmacol Sci* 38(3): 226-256.
11. Mardamshina M, Geiger T (2017) Next-Generation Proteomics and Its Application to Clinical Breast Cancer Research. *Am J Pathol* 187(10): 2175-2184.