

Genes, proteins and metabolites

Opinion

Biology has been an inexact science. Traditionally, it has been a phenomenological science but the situation has not progressed much in spite of tremendous advancement made in the area of molecular biology and other related areas. Much of the modern advances have been contributed by non-traditional biologists, scientists basically trained originally in other disciplines.

We believed that if we could find a proper (highly specific) inhibitor to a specific protein or enzyme, we can perhaps conquer all (well, almost all) the ills the human body suffers. We have made progress but the hope has not been realized. Old drugs do not work in many cases and newer drugs take longer and longer to come to the people. It is still not clear how bacteria develop drug resistance; is it in their genes? Today we know that all the ills of the human body are really not due to external agents (bacteria and viruses).

We had some idea of genes and genetics when Mendel proposed the three famous laws more than 150 years ago. The structure of DNA double helix was determined more than 50 years back. We have already sequenced the more than 3 billion base pairs present in less than two dozen chromosomes some ten years back. We still cannot read the DNA sequence like a novel but we have made considerable progress. There are about 30,000 genes and each gene can make more than one protein. We know a lot about how the genes work and how they are regulated. Our understanding is not complete but it is still a lot.

That is what makes our understanding qualitative in nature. We know a lot about proteins, we know (many of) their structures and functions and we also know how they are regulated. A large number of proteins are enzymes and we know a lot about the reactions they catalyze, the molecules, the substrates and products, the metabolites, in short, the network- the famous metabolic pathway. Most of the metabolic pathway has been worked out painstakingly by biologists over a long period of time. The KEGG database (for the pathway) and the Brenda database (for the enzymes) provide invaluable information for this purpose. What is lacking, then?

Even though we know a lot, it appears that we are still puzzled. That is, in simple words, our lack of understanding of the dynamics of the system. We mean by dynamics, the response of the system under perturbation. How does the system respond when one of the components change, say the glucose concentration, by a relatively large amount? We mean, by the expression, 'relatively large' that the component metabolite in question reaches the extreme (end values for normal subject ranges) concentration values (with other metabolites staying in their preferred normal range). The question, apparently simple, is not easy to answer. We often get qualitative and somewhat evasive, answers that are rarely convincing.

Biochemical processes follow a relatively well understood hierarchical system; the genes, part of the DNA, are relatively inactive except during the cell division. The proteome, however, is gene product and is probably dynamically controlled. In that sense parts (small parts, mind you) of the genome is probably active most of the time during the cell's life cycle. The metabolome, on the other

Volume 5 Issue 5 - 2017

Chanchal K Mitra

Professor of Biochemistry, University of Hyderabad, India

Correspondence: Chanchal K Mitra, Professor of Biochemistry University of Hyderabad, India, Tel +91 40 2313 4668, Fax +91 40 2301 0120, Email c_mitra@yahoo.com

Received: February 21, 2017 | **Published:** April 28, 2017

hand, is omnipresent and perhaps undergoes regular temporal and spatial variations (well, cells too have compartments) in response to various cellular processes. Most common disturbances within the metabolome is most likely due to various external factors (other factors, for example, lack of vitamins in food intake) or errors in the pathway (interprocess inefficiencies due to, for example, defective, or reduced activities of, or missing, enzymes). Some of the problems are genetic and many have been identified with a defect in the gene as a result of the genome sequencing. Most of the genetic diseases are often inherited and are relatively rare.

It is clearly established that the three networks, the network of the genome, the network of the proteome and the metabolome are intimately coupled but function at different levels. The gene network is often triggered by proteins, produces proteins (sometimes RNAs) and is rather slow in response. The protein network is on the other hand is rarely triggered by genes but is regulated by other metabolites (including proteins) and is relatively fast in response. There is a set of basic essential proteins that are omnipresent but many proteins are produced from their genes on demand in response to several signals. In contrast, the metabolome is fast acting, strongly interacting with their parents (the metabolic pathway) and have their concentrations tightly regulated by, most commonly, feedback regulation of the enzymes. Because they are fastest in response, the dynamics within the metabolome is very important because the system can jump to alternate stable points.

Metabolite concentrations in a cell are never a constant and can vary over some ranges. The cell needs a constant (relatively speaking) source of energy to stay at the steady state that is far from equilibrium. Part of the energy is used for maintenance (using fresh or recycled materials) of the system and part of the energy is used to just stay at the same place. The processes involved are non-linear and a perturbation can land the system at another stable but different from the initial, state. Because of the difference in the scale (time scale of response), the three networks can be studied independently and any inter network interactions can be treated and studied as small perturbations.

A casual look at the metabolic pathways shows the presence of several causal loops that serve multiple functions. They confer robustness to the system but much more importantly, they clearly and convincingly suggest that the members of the cyclic pathways must be present together. Apart from the constant supply of energy the cell needs, it also depends on several important external signals. Some of these signals are received via transduction while some are received directly in the form of molecules. Some of these signals too have the

potential to shift the present steady state of the cellular environment to a new one, distinct from the current one.

The metabolic pathways are the static picture of the cell metabolome over the life span of the cell. Any given snapshot of the cell with respect to its metabolome spectrum and composition will necessarily be a function of time- the life of a cell from birth to death. It is important to simulate the system for a better understanding but the tools available today (say the SBML) appear not up to the task. The basic problem appears to be very large size of the pathways database and the incomplete information about the enzymes and metabolites. But without a model, even if it is semi quantitative, we cannot progress far.

Complex interactions are common not only in biology but also in social sciences, politics, economics and other branches. In making a dynamic model, we do not need highly accurate data but we need a reliable set of interaction parameters. Such a model may not predict exact behavior but we shall be more than happy to study the approximate pattern of behavior. Once we are confident with the system we can couple the metabolome network with the proteome network and finally the genome network. The overall result may be

unreliable (in some sense) in terms of predictability but we should not forget the real system is also rather unpredictable. To get better dynamics, we need better parameters.

In some sense the genome network is more tightly coupled with the proteome network but at this time our understanding of the genome is far from adequate. We need better understanding of the regulation and the activation (dynamic selection) process. We just know that transcription factors regulate the gene expressions but we do not know how the transcription factors are themselves regulated. They (TFs) are gene products and they are perhaps further regulated by miRNAs. Much of these confusions will get cleared once we learn to simulate the network and compare and contrast against realistic systems.

But the question remain, where are we in the knowledge path?

Acknowledgements

None.

Conflict of interest

The author declares no conflict of interest.