Int Microbiol (2002) 5: 51–52
DOI 10.1007/s10123-002-0065-0

EDITORIAL

M. Schaechter · J.L. Ingraham

What limits genomics, proteomics, transcriptomics?

Published online: 29 May 2002

Studies on gene expression at a global scale are being carried out with breathtaking speed. Entire genome sequences, levels of mRNAs and catalogues of all the detectable proteins in a cell are subjects of almost daily publications. As with other extraordinary methodological breakthroughs, the current ones promise nothing short of a paradigm shift. Biology, accordingly, may become largely a computational science, an opinion viewed by some with mixed emotions. Perhaps most readers of this journal agree that work along more traditional lines must continue to receive its fair share of emphasis and that the computer and the bench must dwell under the same tent.

Here, we would like to pose some questions that may not be answered by global studies alone and that might require both traditional and novel approaches. For many of these questions, the answers will be slow to come, a delay that may test the patience of investigators. That is the way of science.

Here are some examples of questions that appear to us to surpass the power of present-day global studies:

Protein activity

How many of the protein molecules in a cell are active? What are the "real times" of biochemical reactions in living cells? Does the law of mass action alone determine the frequency of collision of small numbers of molecules with their target ? Analogous questions can be posed for mRNAs. The answers are compounded greatly by differences in the stability between individual molecular species. Thus, we are reminded, measuring the total number of molecules present in a cell does not tell us how many are actually made, leave alone how many are active.

Cellular locations

Where in the cell are biochemical transactions carried out? Increasingly, the application of fluorescence and other kinds of microscopy tell us the location of specific proteins and DNA sequences in cells. However, are the foci of localized macromolecules revealed by these methods the actual site of macromolecular activity or are these aggregates molecular reservoirs? Answering such questions requires quantitation of molecular *activity* and discrimination between molecular species, both of which are difficult to achieve. Progress in single-molecule chemistry may well help here. In addition, there is much to learn about the internal milieu of the cell. In the apt words of Mahadevan and Matsudaira "In this soft, wet, and dynamic world, structural features are dominated by filamentous and membranous objects, a constant reminder that all events at this level are mediated by interfacial interactions" [Mahadevan L, Matsudaira P (2000) Motility powered by supermolecular springs and ratchets. Science 288:95–100].

Gene expression

The aim of many global studies is to determine the effects of specific imposed conditions on gene expression or protein levels. A grand example is the work done on gene expression in yeast (see http://genome-www.stanford.edu/Saccharomyces/). The level of significance of changes is often set arbitrarily, typically a factor of 2 between two conditions. Yet, one can imagine situations in which a two-fold change in a protein or an mRNA may have little physiological consequence, and others where a 20% change may be critical. The trouble here is

M. Schaechter (✉)
Department of Biology, San Diego State University, California, USA
E-mail: mschaech@sunstroke.sdsu.edu

J.L. Ingraham
Department of Microbiology.
University of California-Davies, USA

that bioinformatics, in its raw versions, demands a "one size fit all" criterion of significance. With increased sophistication and a measure of diffidence, this problem may become tractable. Of course, a change is comprehensible and reproducible only if the starting point is clearly defined. Tragically, many global studies are done on cultures grown in complex media for an unspecified period. Under such conditions, the composition of the medium and therefore, the physiological state of the cells is unknown at the starting point. Is it too much to ask for standard experimental conditions?

The basis of changes in gene expression

Some of the changes in gene expression or protein levels determined en masse are specific to the conditions under study, others not. Some are due to global responses; others are specifically elicited by the condition imposed. Collecting vast amounts of data may well help sort out which is which. Furthermore, why are there so many circuits to regulate gene expression? In bacteria, the repertoire of control mechanisms is formidable, but the real life importance of each is elusive. Why so many, and when and where do they function?

Annotation

For many reasons, annotation, the attribution of function to genes is a complicated and vexing business. For example, many, perhaps most proteins are multifunctional and participate in more than one biochemical activity, often as part of protein complexes. Moreover, sequential metabolic steps catalyzed by a single multifunctional enzyme in one organism are often catalyzed by several unifunctional enzymes in others. The tryptophan biosynthesis pathway in bacteria is a worthy example. Sequence similarity between enzymes with different function is an added complication. Programs settle for a close fit without searching for a closer one. The startlingly excess of malic over lactic dehydrogenase resulting from computational annotation illustrates the point. In addition, backups and redundancies commonly occur. Which of the functionally cognate molecules works under what circumstances? Annotation (as with the blind men and the elephant) is giving a single name to an entity that has several.

The morphome

Morphome, the complete structure of the cell, an ever more popular word in computational biology, is becoming its Holy Grail. But will we, as the word's currency implies, ever be able to deduce a cell's totality from its genome sequence? Franklin Harold (and we too) think not [Harold FM (2001) *The way of the cell: molecules, organisms and the order of life*. Oxford University Press, Oxford] One problem is the genome's dependence on the cell. A simple example is the synthesis of ATP. Even knowing all gene products and their function, the cleverest computer program would consider ATP synthesis impossible unless it was told ADP was already present in the cell. Many of these questions will have solutions, some to arise sooner, some later. What lies beyond? In narrow terms, global studies are the "what," not the "where" or "when." Currently, global studies, however sophisticated, are essentially enumerative. But, coupled with "one-gene-at-a-time" experimentation, global studies will paint for us a grand picture of when genes are turned on and off. Let us anticipate the day when coupled studies will have been carried out as extensively as is reasonable and with a staggering degree of sophistication. We will be able to make predictions about the structure of proteins and figure out the rules that govern their work. We will then know a lot about phenomena such as the cell cycle, developmental processes, responses to mild and not so mild environmental stresses, and probably others. And that's not all. Although the horizon of our questions is likely to recede, the acquisition of knowledge is incremental and breakthrough feeds on breakthrough.

Beyond these considerations, however, remain the overarching questions of ecology and evolution. Happily, the study of microbes in natural environments is becoming quite popular. There is increased realization that these are not isolated fields but that they are central to gaining an understanding of how the pieces of the puzzle fit together. With novel approaches, we will be able to ask deeper and deeper questions about what it is that organisms actually do in The Real World.

Moselio Schaechter is Distinguished Professor of Molecular Biology and Microbiology, Emeritus, Tufts University School of Medicine, Boston, Massachusetts, and past president of the American Society for Microbiology (1986). He lives now in San Diego, California. He coauthored *Mechanisms of Microbial Disease, Physiology of the Bacterial Cell*, and Escherichia coli *and* Salmonella. *Cellular and Molecular Biology*, among other books.

John L. Ingraham is Emeritus Professor, Section of Microbiology, University of California, Davis, and past president of the American Society for Microbiology (1993). He has coauthored *The Microbial World* (since its 4th edition), *The Growth of the Bacterial Cell*, *The Physiology of the Bacterial Cell* and *Introduction to Microbiology* (the latter with his daughter Catherine Ingraham Vigran). He was Associate Editor of *Annual Review of Microbiology*, North American editor of *Archives of Microbiology*, and Editor-in-Chief of *Microbiological Reviews* (formerly *Bacteriological Reviews* and now *Microbiological and Molecular Biological Reviews*).